

BUNDESREPUBLIK DEUTSCHLAND

REC'D 13 OCT 2004

WIPO

PCT

**Prioritätsbescheinigung über die Einreichung
einer Patentanmeldung****Aktenzeichen:**

10 2004 022 065.4

Anmeldetag:

5. Mai 2004

Anmelder/Inhaber:

Degussa AG, 40474 Düsseldorf/DE

Bezeichnung:

Screeningverfahren für Hydantoinrazemasen

Priorität:

6. Juni 2003 DE 103 26 109.5

IPC:

C 12 N, C 12 Q, C 07 H

**PRIORITY
DOCUMENT**
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

BEST AVAILABLE COPY

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 27. September 2004
Deutsches Patent- und Markenamt
Der Präsident
Im Auftrag

Dzierzon

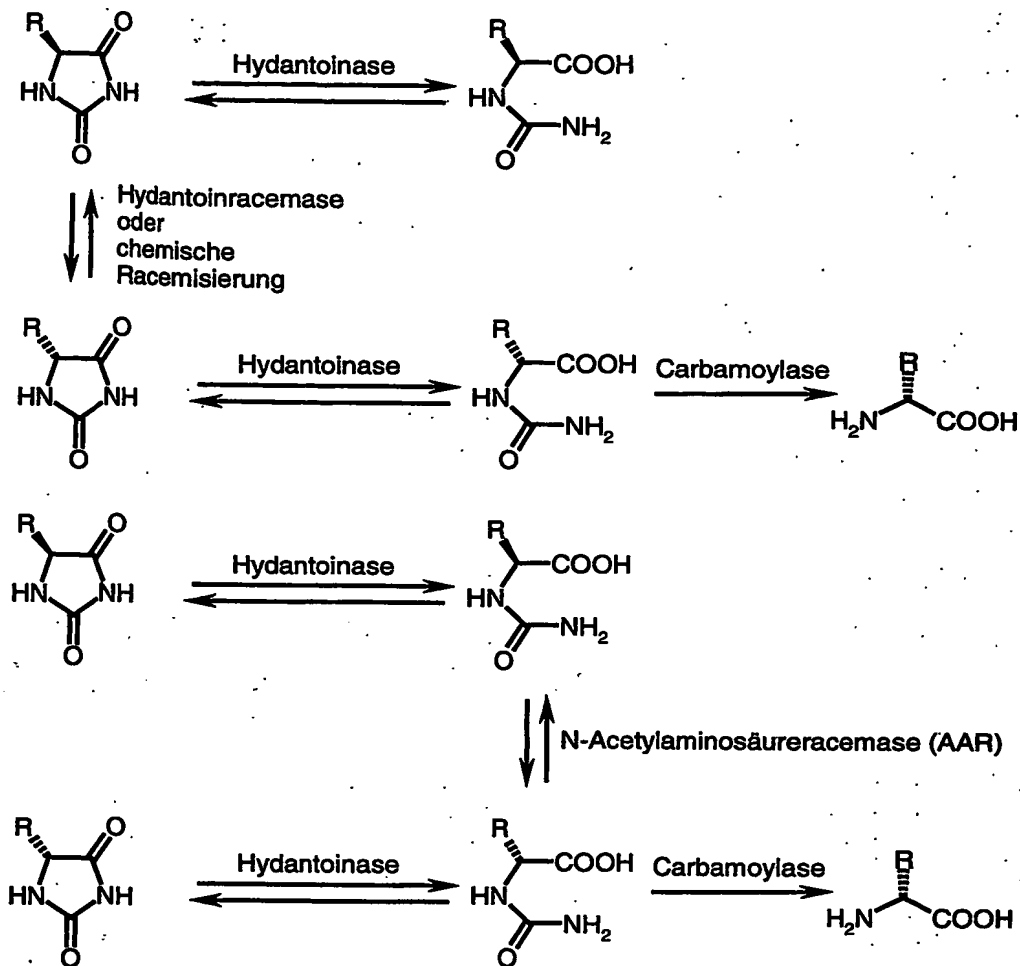
Screeningverfahren für Hydantoinrazemasen

Die vorliegende Erfindung ist auf ein Screeningverfahren zur Detektion verbesserter Hydantoinrazemasen, neue Hydantoinrazemasen selbst und deren Verwendung zur Herstellung von N-Carbamoyl-Aminosäuren gerichtet.

Diese optisch aktiven Verbindungen sind in der organischen Synthese zur Herstellung von z.B. bioaktiven Wirkstoffen häufig eingesetzte Verbindungen. Sie kommen auch in chiralen Auxiliaren z.B. in Form der Aminoalkohole (Evans-Reagenzien) vor.

Die enzymatische Hydrolyse von 5-substituierten Hydantoinen zu N-Carbamoyl-Aminosäuren und deren Weiterreaktion zu den entsprechenden enantiomerenangereicherten Aminosäuren ist eine Standardmethode in der organischen Chemie ("Enzyme Catalysis in Organic Synthesis", Eds.: Drauz, Waldmann, VCH, 1st and 2nd Ed.). Die Enantiodifferenzierung kann dabei entweder auf der Stufe der Hydantoinhydrolyse durch Hydantoinasen erfolgen oder aber wahlweise bei der Spaltung der N-Carbamoyl-Aminosäuren mittels enantioselektiver Carbamoylasen. Da die Enzyme nur jeweils eine optische Antipode der entsprechenden Verbindung umsetzen, wird versucht, die andere im Gemisch (in-situ) zu razemisieren, um den vollständigen Umsatz des razemisch leicht herstellbaren Hydantoins in die korrespondierende enantiomerenangereicherte Aminosäure zu gewährleisten. Die Razemisierung kann dabei entweder auf der Stufe der Hydantoine mittels chemischer (Base, Säure, erhöhte Temp.) oder enzymatischer Verfahren erfolgen oder aber auf der Stufe der N-Carbamoyl-Aminosäuren mittels z.B. Acetylaminosäurerazemasen (DE10050124) vonstatten gehen. Letztere Variante funktioniert erfolgreich naturgemäß nur bei Einsatz von enantioselektiven Carbamoylasen. Das nachfolgende Schema veranschaulicht diesen Sachverhalt.

Schema 1:



5

Für aromatische Substrate ist die Geschwindigkeit der chemischen Razemisierung der Hydantoine, wie in Tabelle 1 gezeigt, ausreichend hoch, um hohe Raum-Zeit-Ausbeuten für die Herstellung von Aminosäuren nach dem

- 10 Hydantoinaseverfahren zu gewährleisten. Für aliphatische Hydantoine wie Isobutyl-, Methyl- und Isopropylhydantoin stellt die Razemisierung jedoch einen erheblichen Engpass bei der Synthese aliphatischer Aminosäuren dar.

Tabelle 1: Razemisierungskonstanten von Hydantoinen bei 40°C, pH 8.5 bestimmt durch Anfangsraten gem. einer Reaktion erster Ordnung ($-k_{rac} = \ln([a]/[a_0])$) aus: Hydrolysis and Formation of Hydantoins (Chpt. B 2.4). Sylatk, C. and Pietzsch, M. In: Enzyme catalysis in organic synthesis (Eds.: K. Drauz & H. Waldmann), VCH, 1st and 2nd Ed.).

5'-substituent	k_{rac} (h ⁻¹)	$t_{1/2}$ (h)
Phenyl	2.59	0.27
Methylthioethyl	0.12	5.82
Isobutyl	0.032	21.42
Methyl	0.02	33.98
Isopropyl	0.012	55.90

Dieses Problem zeigt sich beispielsweise bei der in EP759475 beschriebenen Herstellung von enantiomerenangereichertem tert-Butylhydantoin mittels des Hydantoinaseverfahrens. Hier wurden zur vollständigen Umsetzung von 32mM tert.-Butylhydantoin mit 1,5kU R-Hydantoinase 8 Tage bei pH 8,5 und 4 Tage bei pH 9,5 benötigt. Tatsächlich ist die geringe Raum-Zeit-Ausbeute durch die nur langsame chemische Razemisierung von tert-Butylhydantoin ($k_{rac} = 0.009\text{h}^{-1}$ bei 50°C und pH 8.5) bedingt.

Aus dem Stand der Technik sind Hydantoinrazemasen aus Mikroorganismen der Gattung *Pseudomonas*, *Mikrobacterium*, *Agrobacterium* und *Arthrobacter* bekannt (Lit.: JP04271784; EP1188826; Cloning and characterization of genes from *Agrobacterium* sp. IP I-671 involved in hydantoin degradation. Hils, M.; Muench, P.; Altenbuchner, J.; Sylatk, C.; Mattes, R. Applied Microbiology and Biotechnology (2001), 57(5-6), 680-688; A new razemase for 5-monosubstituted hydantoins. Pietzsch, Markus;

Syldatk, Christoph; Wagner, Fritz. Ann. N. Y. Acad. Sci. (1992), 672 (Enzyme Engineering XI), 478-83. Lickefett, Holger; Krohn, Karsten; Koenig, Wilfried A.; Gehrcke, Barbel; Syldatk, Christoph. Tetrahedron: Asymmetry (1993), 4(6), 1129-35; Purification and characterization of the hydantoin razemase of *Pseudomonas* sp. strain NS671 expressed in *Escherichia coli*. Watabe, Ken; Ishikawa, Takahiro; Mukohara, Yukuo; Nakamura, Hiroaki. J. Bacteriol. (1992), 174(24), 7989-95).

10 Von den Hydantoinrazemasen aus *Arthrobacter aurescens* DSM 3745, *Pseudomonas* sp. NS671 und *Microbacterium liquefaciens* ist bekannt, dass diese Enzyme aliphatische Hydantoine wie beispielsweise Isopropylhydantoin oder Isobutylhydantoin nur schwach razemisieren. Darüber hinaus weiß man, dass die
15 Hydantoinrazemasen aus *Arthrobacter aurescens* DSM 3747 aromatische Hydantoine wie Indolylmethylhydantoin oder Benzylhydantoin bevorzugt, wohingegen aliphatische Hydantoine wie Methylthioethylhydantoin vergleichsweise schwach oder im Fall von Isopropylhydantoin überhaupt nicht
20 umgesetzt werden (A new razemase for 5-monosubstituted hydantoins. Pietzsch, Markus; Syldatk, Christoph; Wagner, Fritz. Ann. N. Y. Acad. Sci. (1992), 672 (Enzyme Engineering XI), 478-83.)).

Die niedrige Aktivität von Hydantoinrazemasen begrenzt daher häufig das wirtschaftliche Potential dieser Route.

Um in geeigneter Zeit möglichst viele Hydantoinrazemasen auf ihr Potential zur Razemisierung von aliphatischen Hydantoine prüfen zu können, lag die Aufgabe der vorliegenden Erfindung unter anderem in der Angabe eines
30 geeigneten Screeningverfahrens für Hydantoinrazemasen. Darüber hinaus sollte das erfindungsgemäße Screeningverfahren als Bestandteil für ein Mutageneseverfahren zur Gewinnung neuer und verbesserter Hydantoinrazemasen einsetzbar sein. Ebenfalls Aufgabe der
35 vorliegenden Erfindung war die Angabe neuer

Hydantoinrazemasen, die den Hydantoinrazemasen des Standes der Technik zumindest in Selektivität und/oder Aktivität und/oder Stabilität überlegen sind.

Diese Aufgabe wird anspruchsgemäß gelöst. Anspruch 1

- 5 bezieht sich auf ein Screeningverfahren für Hydantoinrazemasen. Unteransprüche 2 bis 4 zeigen vorteilhafte Ausführungsformen des Screeningverfahrens auf. Anspruch 5 beschäftigt sich mit einem Mutageneseverfahren zur Herstellung neuer Hydantoinrazemasen unter Anwendung des erfindungsgemäßen Screeningverfahrens. Ansprüche 6 bis 10 beziehen sich auf neue Hydantoinrazemasen sowie die sie codierenden Nukleinsäuresequenzen und deren Verwendung. Ansprüche 12 bis 14 richten sich auf Vehikel, welche die erfindungsgemäßen Hydantoinrazemasen aufweisen, bzw. 15 spezielle Primer für deren Herstellung.

Dadurch, dass man ein Screeningverfahren für Hydantoinrazemasen angibt, bei dem man

- a) eine enantioselektive Hydantoinase und
b) die zu prüfende Hydantoinrazemase, welche eine 20 verglichen mit der Hydantoinase unter a) langsamere Umsetzungsrate aufweist, auf
c) ein chirales Hydantoin einwirken lässt, welches in zur Selektivität der Hydantoinase entgegengesetzter enantiomerenangereicherter Form 25 eingesetzt wird, und
d) die resultierende N-Carbamoyl-Aminosäure oder die freigesetzten Protonen zeitabhängig detektiert, gelangt man überraschend einfach und dennoch vorteilhaft zu einer Möglichkeit, viele Hydantoinrazemasen in kurzer Zeit auf 30 ihre Fähigkeit hin zu überprüfen, in verbesserter Weise Hydantoine razemisieren zu können.

Durch Einsatz eines L-Enantiomers eines 5'-monosubstituierten Hydantoins und Verwendung einer D-selektiven Hydantoinase, welche aufgrund ihrer 35 Enantioselektivität bevorzugt das entstehende D-Enantiomer

- des Hydantoins schnell hydrolysiert, kann durch die Bildung der N-Carbamoyl-D-Aminosäure oder freiwerdende Protonen die Razemisierungsgeschwindigkeit und damit die Aktivität der Hydantoinrazemase auf einfache Weise gemessen werden. Die
- 5 Quantifizierung der N-Carbamoyl-Aminosäure kann dabei durch dem Fachmann bekannte Methoden wie beispielsweise HPLC oder colorimetrische Methoden erfolgen. Die Quantifizierung über Protonen kann auf einfache Weise über pH Indikatoren, bevorzugt Cresol Rot, erfolgen. Es sei darauf hingewiesen,
- 10 dass in dem Verfahren sowohl D- als auch L-Enantiomere von Hydantoinen mit unterschiedlichen ggf. aliphatischen 5'-Substituenten eingesetzt werden können. Beim Einsatz der D-Hydantoine sind dementsprechenden L-selektive Hydantoinasen im Screeningverfahren einzusetzen.
- 15 Im erfindungsgemäßen Verfahren eingesetzt werden vorteilhaft aliphatische in 5'-Stellung substituierte Hydantoine. Unter aliphatisch substituierten Hydantoinen wird in diesem Zusammenhang ein System verstanden, welches in 5'-Stellung an dem Hydantoinheterozyklus einen Rest
- 20 aufweist, der über ein C-Atom mit sp^3 -Hybridisierung an den Heterozyklus gebunden ist. Bevorzugte 5'-Substituenten sind dabei Methyl, Ethyl, Butyl, Propyl, tertiär-Butyl, Isopropyl und Isobutyl. Ganz besonders bevorzugt ist Ethyl-Hydantoin.
- 25 Als Hydantoinasen können sämtliche in der Literatur bekannten Hydantoinasen eingesetzt werden, welche das über die Hydantoinrazemase gebildete Enantiomer des Hydantoins enantioselektiv hydrolysieren, wobei diese Hydrolyse schneller als die Razemisierungsgeschwindigkeit sein muss.
- 30 Bevorzugte Hydantoinasen sind dabei die kommerziellen Hydantoinasen 1 & 2 von Roche, die Hydantoinasen der Gattungen *Agrobacterium*, *Arthrobacter*, *Bacillus*, *Pseudomonas*, *Flavobacterium*, *Pasteurella*, *Microbacterium*, *Vigna*, *Ochrobactrum*, *Methanococcus*, *Burkholderia* und
- 35 *Streptomyces*. (Hils, M.; Muench, P.; Altenbuchner, J.;

- Syldatk, C.; Mattes, R. Cloning and characterization of genes from *Agrobacterium* sp. IP I-671 involved in hydantoin degradation. *Applied Microbiology and Biotechnology* (2001), 57(5-6), 680-688. Soong, C.-L.; Ogawa, J.; Shimizu, S. Cyclic ureide and imide metabolism in microorganisms producing a D-hydantoinase useful for D-amino acid production. *Journal of Molecular Catalysis B: Enzymatic* (2001), 12(1-6), 61-70. Wiese, Anja; Wilms, Burkhard; Syldatk, Christoph; Mattes, Ralf; Altenbuchner, Josef. Cloning, nucleotide sequence and expression of a hydantoinase and carbamoylase gene from *Arthrobacter aureus* DSM 3745 in *Escherichia coli* and comparison with the corresponding genes from *Arthrobacter aureus* DSM 3747. *Applied Microbiology and Biotechnology* (2001), 55(6), 750-757. Yin, Bang-Ding; Chen, Yi-Chuan; Lin, Sung-Chyr; Hsu, Wen-Hwei. Production of D-amino acid precursors with permeabilized recombinant *Escherichia coli* with D-hydantoinase activity. *Process Biochemistry* (Oxford) (2000), 35(9), 915-921. Park, Joo-Ho; Kim, Geun-Joong; Lee, Seung-Goo; Lee, Dong-Cheol; Kim, Hak-Sung. Purification and characterization of thermostable D-hydantoinase from *Bacillus thermocatenulatus* GH-2. *Applied Biochemistry and Biotechnology* (1999), 81(1), 53-65; Pozo, C.; Rodelas, B.; de la Escalera, S.; Gonzalez-Lopez, J. D,L-Hydantoinase activity of an *Ochrobactrum anthropi* strain. *Journal of Applied Microbiology* (2002), 92(6), 1028-1034; Chung, Ji Hyung; Back, Jung Ho; Lim, Jae-Hwan; Park, Young In; Han, Ye Sun. Thermostable hydantoinase from a hyperthermophilic archaeon, *Methanococcus jannaschii*. *Enzyme and Microbial Technology* (2002), 30(7), 867-874; Xu, Zhen; Jiang, Weihong; Jiao, Ruishen; Yang, Yunliu. Cloning, sequencing and high expression in *Escherichia coli* of D-hydantoinase gene from *Burkholderia pickettii*. *Shengwu Gongcheng Xuebao* (2002), 18(2), 149-154; Las Heras-Vazquez, Francisco Javier; Martinez-Rodriguez, Sergio; Mingorance-Cazorla, Lydia; Clemente-Jimenez, Josefa Maria; Rodriguez-Vico, Felipe.

Overexpression and characterization of hydantoin racemase from *Agrobacterium tumefaciens* C58. Biochemical and Biophysical Research Communications (2003), 303(2), 541-547; DE 3535987; EP 1275723; US 6087136; WO 0281626; US 5 2002045238; DE 4328829; WO 9400577; WO 9321336; JP 04325093; NL 9001680; JP 2003024074; WO 0272841; WO 0119982; WO 9620275).

10 Ganz besonders bevorzugt ist die Verwendung der Hydantoinase aus *Arthrobacter crystallopoietes*, insbesondere der aus DSM 20117.

Wie schon angedeutet sollte die Umsetzungsgeschwindigkeit der Hydantoinase die der Razemase übertreffen. Vorzugsweise liegt das Verhältnis der Geschwindigkeitskonstanten der Hydantoinase zur Hydantoinrazemase ($k_{\text{Hyd}}/k_{\text{Raz}}$) bei >2 ,
15 besonders bevorzugt bei >10 und ganz besonders bevorzugt bei >50 .

Gegenstand der Erfindung ist ebenfalls ein Verfahren zur Herstellung von verbesserten Hydantoinrazemasen, welches sich dadurch auszeichnet, dass man

- 20 a) die Nukleinsäuresequenz codierend für die Hydantoinrazemase einer Mutagenese unterwirft,
b) die aus a) erhältlichen Nukleinsäuresequenzen in einen geeigneten Vektor kloniert und diesen in ein geeignetes Expressionssystem transferiert und
25 c) die gebildeten Hydantoinrazemasen mit verbesserter Aktivität und/oder Selektivität und/oder Stabilität mittels eines erfindungsgemäßen Screeningverfahrens detektiert und isoliert.

Als Ausgangsgene für die Mutagenese der Hydantoinrazemasen
30 können sämtliche bekannten und in der angeführten Literatur erwähnten Hydantoinrazemasegene dienen. Bevorzugt sind dabei die Hydantoinrazemasegene von *Arthrobacter*, *Pseudomonas*, *Agrobacterium* und *Micrococcus* (Wiese A; Pietzsch M; Syldatk C; Mattes R; Altenbuchner J Hydantoin
35 racemase from *Arthrobacter aurescens* DSM 3747: heterologous

expression, purification and characterization. JOURNAL OF BIOTECHNOLOGY (2000 Jul 14), 80(3), 217-30; Watabe K; Ishikawa T; Mukohara Y; Nakamura H Purification and characterization of the hydantoin racemase of *Pseudomonas* sp. strain NS671 expressed in *Escherichia coli*. JOURNAL OF BACTERIOLOGY (1992 Dec), 174(24), 7989-95; Las Heras-Vazquez, Francisco Javier; Martinez-Rodriguez, Sergio; Mingorance-Cazorla, Lydia; Clemente-Jimenez, Josefa Maria; Rodriguez-Vico, Felipe. Overexpression and
5 characterization of hydantoin racemase from *Agrobacterium tumefaciens* C58. Biochemical and Biophysical Research Communications (2003), 303(2), 541-547; EP 1188826). Ganz besonders bevorzugt ist das Hydantoinrazemasegen aus *Arthrobacter aurescens* welches für die Proteinsequenz in
10 Seq.ID.Nr. 2 codiert.

Zur Mutagenese der Hydantoinrazemase können sämtliche in der Literatur bekannten Methoden wie beispielsweise Zufallsmutagenese, Sättigungsmutagenes, Kassetten-Mutagenese oder Rekombinationsmethoden verwendet werden
15 (May, Oliver; Voigt, Christopher A.; Arnold, Frances H. Enzyme engineering by directed evolution. Enzyme Catalysis in Organic Synthesis (2nd Edition) (2002), 1 95-138; Bio/Technology 1991, 9, 1073-1077; Horwitz, M. und Loeb, L., Promoters Selected From Random DNA-Sequences, Proc Natl Acad Sci USA 83, 1986, 7405-7409; Dube, D. und L. Loeb, Mutants Generated By The Insertion Of Random
20 Oligonucleotides Into The Active-Site Of The Beta-Lactamase Gene, Biochemistry 1989, 28, 5703-5707; Stemmer, P.C., Rapid evolution of a protein in vitro by DNA shuffling, Nature 1994, 370, 389-391 und Stemmer, P.C., DNA shuffling
30 by random fragmentation and reassembly: In vitro recombination for molecular evolution. Proc Natl Acad Sci USA 91, 1994, 10747-10751).

Die Klonierung und Expression kann wie in der weiter unten
35 angegebenen Literatur durchgeführt werden. Das Verfahren kann mehrmals hintereinander ggf. mit wechselnden Mutagenesestrategien durchgeführt werden.

Gegenstand der Erfindung sind ebenfalls rec-Polypeptide oder die diese codierende Nukleinsäuresequenzen, welche nach dem eben genannten Mutageneseverfahren erhältlich sind.

- 5 Ebenso ein Aspekt der Erfindung ist die Verwendung der so hergestellten Polypeptide zur Herstellung von chiralen enantiomerenangereicherten N-Carbamoyl-Aminosäuren oder Aminosäuren. Die erfindungsgemäß hergestellten Nukleinsäuresequenzen können zur Herstellung von
- 10 Ganzzellkatalysatoren dienen.

- Einen Teil der vorliegenden Erfindung bilden auch Hydantoinrazemasen, welche in Position 79 einen Aminosäureaustausch mit einer Aminosäure ausgewählt aus der Gruppe bestehend aus A, R, N, D, C, Q, E, H, I, L, K, M, F,
- 15 P, S, T, Y oder V aufweisen. Interessant ist, dass die Aminosäuren, welche diese Position umgeben, für viele Hydantoinrazemasen vollständig konserviert sind. Die Konsensussequenz lautet: FX_1DX_2GL (Seq.ID.Nr. 1), wobei X_2 P oder T darstellt und X_1 W oder G darstellt. Bevorzugte
- 20 Mutanten weisen daher die oben genannte Konsensussequenz auf, wobei X_1 vorzugsweise eine Aminosäure ausgewählt aus der Gruppe bestehend aus A, R, N, D, C, Q, E, H, I, L, K, M, F, P, S, T, Y oder V darstellt. X_1 entspricht dabei der Position 79. Bevorzugte Mutanten sind in Tabelle 2
- 25 dargestellt.

Tabelle 2:

Mutanten Name	Mutation (codon)	Mutation X ₁ (Aminosäure)	Aktivitäts-änderung	Seq.ID Nr.
3CH11	GGG -> GAG	G79E	2	5
1BG7	GGG -> AGG	G79R	2	3
BB5	GGG -> TTG	G79L	4	9
AE3	GGG -> CAG	G79Q	4	7

Weitere äußerst vorteilhafte Kombinationen von X₁ und X₂ Hydantoinrazemasen sind in folgender Tabelle 3 aufgeführt.

Tabelle 3: Vorteilhafte Kombinationen von X₁ und X₂ in dem Konsensusmotiv FX₁DX₂GL

X ₁	L	E	Q	R	L	E	Q	R
X ₂	P	P	P	P	T	T	T	T

Von besonderem Vorteil ist es, wenn die Hydantoinrazemasen die oben angegebene Konsensusregion und zusätzlich eine Homologie von >40% zur Hydantoinrazemase aus DSM 20117 aufweisen.

Weiterhin Gegenstand der Erfindung sind isolierte Nukleinsäuresequenzen codierend für eine Hydantoinrazemase ausgewählt aus der Gruppe:

- einer Nukleinsäuresequenz codierend für eine erfindungsgemäße Hydantoinrazemase,
- einer Nukleinsäuresequenz, die unter stringenten Bedingungen mit der Nukleinsäuresequenz codierend für

eine erfindungsgemäße Hydantoinrazemase oder der dazu komplementären Sequenz hybridisiert,

c) einer Nukleinsäuresequenz gemäß den Seq.ID.Nr. 3, 5, 7 oder 9 oder solchen mit einer Homologie von > 80% zu diesen,

d) einer Nukleinsäuresequenz aufweisend 15 aufeinanderfolgende Nukleotide der Sequenzen Seq.ID.Nr. 3, 5, 7 oder 9.

In Bezug auf Punkt d) ist es bevorzugt, wenn die erfindungsgemäße Nukleinsäuresequenz 20, mehr bevorzugt 25, weiter bevorzugt 30, 31, 32, 33, 34 und äußerst bevorzugt mehr als 34 identische konsekutive Nukleinsäuren der Sequenzen Seq.ID.Nr. 3, 5, 7 oder 9 aufweist.

Wie gesagt sind von der Erfindung auch

Nukleinsäuresequenzen mitumfasst, welche unter stringenten Bedingungen mit den erfindungsgemäßen einzelsträngigen Nukleinsäuresequenzen oder deren komplementären einzelsträngigen Nukleinsäuresequenzen hybridisieren (b) oder solche, die sich in Sequenzabschnitten gleichen (d).

Als solche sind z.B. spezielle Gensonden oder die für eine PCR notwendigen Primer anzusehen.

Eine Kopplung von Hydantoinrazemase und Hydantoinase und ggf. Carbamoylase kann dabei durch Zusammengeben der freien bzw. immobilisierten Enzyme erfolgen. Bevorzugt ist jedoch, wenn die Hydantoinase gemeinsam mit der Hydantoinrazemase und/oder der Carbamoylase in der selben Zelle exprimiert wird (Ganzzellkatalysator).

Die erfindungsgemäßen Nukleinsäuresequenzen können daher als Bestandteil eines Gens in analoger Weise wie in DE10234764 und dort zitierter Literatur in einen Ganzzellkatalysator kloniert werden.

Sofern dieser dann auch Gene für eine Hydantoinase und/oder Carbamoylase aufweist, ist er im Stande racemische Hydantoine zur Gänze in enantiomerenangereicherte

Aminosäuren umzuwandeln. Ohne ein kloniertes

Carbamoylasegen stoppt die Reaktion auf der Stufe der N-Carbamoyl-Aminosäuren.

Vorzugsweise wird ein Organismus wie in der DE10155928 genannt als Wirtsorganismus eingesetzt. Der Vorteil eines derartigen Organismus ist die gleichzeitige Expression aller beteiligten Enzyme, womit nur noch ein rec-Organismus für die Gesamtreaktion angezogen werden muss.

Um die Expression der Enzyme im Hinblick auf ihre Umsetzungsgeschwindigkeiten abzustimmen, können die entsprechenden codierenden Nukleinsäuresequenzen in unterschiedliche Plasmide mit unterschiedlichen Kopienzahlen kloniert und/oder unterschiedlich starke Promotoren für eine unterschiedlich starke Expression der Nukleinsäuresequenzen verwendet werden. Bei derart

abgestimmten Enzymsystemen tritt vorteilhafterweise eine Akkumulation einer ggf. inhibierend wirkenden Zwischenverbindung nicht auf und die betrachtete Reaktion kann in einer optimalen Gesamtgeschwindigkeit ablaufen. Dies ist dem Fachmann jedoch hinlänglich bekannt

(Gellissen, G.; Piontek, M.; Dahlems, U.; Jenzelewski, V.; Gavagan, J. W.; DiCosimo, R.; Anton, D. L.; Janowicz, Z. A. (1996), Recombinant *Hansenula polymorpha* as a biocatalyst. Coexpression of the spinach glycolate oxidase (GO) and the *S. cerevisiae* catalase T (CTT1) gene, Appl. Microbiol.

Biotechnol. 46, 46-54; Farwick, M.; London, M.; Dohmen, J.; Dahlems, U.; Gellissen, G.; Strasser, A. W.; DE19920712).

Die Herstellung eines derartigen Ganzzellkatalysators ist dem Fachmann hinlänglich bekannt (Sambrook, J.; Fritsch, E. F. und Maniatis, T. (1989), Molecular cloning: a laboratory

manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York; Balbas, P. und Bolivar, F. (1990), Design and construction of expression plasmid vectors in *E. coli*, Methods Enzymol. 185, 14-37; Rodriguez, R.L. und Denhardt, D. T (eds) (1988), Vectors: a survey of molecular cloning

vectors and their uses, 205-225, Butterworth, Stoneham).

In einer nächsten Ausgestaltung bezieht sich die Erfindung auf Plasmide oder Vektoren aufweisend eine oder mehrere der erfindungsgemäßen Nukleinsäuresequenzen.

Als Plasmide oder Vektoren kommen im Prinzip alle dem

- 5 Fachmann für diesen Zweck zur Verfügung stehenden Ausführungsformen in Frage. Derartige Plasmide und Vektoren können z. B. von Studier und Mitarbeiter (Studier, W. F.; Rosenberg A. H.; Dunn J. J.; Dubendroff J. W.; (1990), Use of the T7 RNA polymerase to direct expression of cloned
- 10 genes, Methods Enzymol. 185, 61-89) oder den Broschüren der Firmen Novagen, Promega, New England Biolabs, Clontech oder Gibco BRL entnommen werden. Weiter bevorzugte Plasmide und Vektoren können gefunden werden in: Glover, D. M. (1985), DNA cloning: a practical approach, Vol. I-III, IRL Press
- 15 Ltd., Oxford; Rodriguez, R.L. und Denhardt, D. T (eds) (1988), Vectors: a survey of molecular cloning vectors and their uses, 179-204, Butterworth, Stoneham; Goeddel, D. V. (1990), Systems for heterologous gene expression, Methods Enzymol. 185, 3-7; Sambrook, J.; Fritsch, E. F. und
- 20 Maniatis, T. (1989), Molecular cloning: a laboratory manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York.

Plasmide, mit denen das die erfindungsgemäße Nukleinsäure aufweisende Genkonstrukt in ganz bevorzugter Weise in den Wirtsorganismus kloniert werden kann, sind Derivate von pUC18 und pUC19 (Roche Biochemicals), pKK-177-3H (Roche Biochemicals), pBTac2 (Roche Biochemicals), pKK223-3 (Amersham Pharmacia Biotech), pKK-233-3 (Stratagene) oder pET (Novagen). Weiterebevorzugte Plasmide sind pBR322

30 (DSM3879), pACYC184 (DSM4439) und pSC101 (DSM6202), welche von der DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Braunschweig, Germany bezogen werden können.

- Als bevorzugt anzusehende Plasmide Gleichfalls ist die
- 35 Erfindung auf Mikroorganismen aufweisend eine oder mehrere erfindungsgemäße Nukleinsäuresequenzen gerichtet. Der Mikroorganismus, in den die die erfindungsgemäßen

- Nukleinsäuresequenzen enthaltenen Plasmide kloniert werden, dient zur Vermehrung und Gewinnung einer ausreichenden Menge des rekombinanten Enzyms. Die Verfahren hierfür sind dem Fachmann wohlbekannt (Sambrook, J.; Fritsch, E. F. und Maniatis, T. (1989), *Molecular cloning: a laboratory manual*, 2nd ed., Cold Spring Harbor Laboratory Press, New York). Als Mikroorganismen können im Prinzip alle dem Fachmann für diesen Zweck in Frage kommenden Organismen wie z.B. Hefen wie *Hansenula polymorpha*, *Pichia sp.*, *Saccharomyces cerevisiae*, Prokaryonten, wie *E. coli*, *Bacillus subtilis* oder Eukaryonten, wie Säugerzellen, Insektenzellen herangezogen werden. Vorzugsweise sind *E. coli*-Stämme für diesen Zweck zu benutzen. Ganz besonders bevorzugt sind: *E. coli* XL1 Blue, W3110, DSM14459 (PCT/US00/08159), NM 522, JM101, JM109, JM105, RR1, DH5 α , TOP 10⁻ oder HB101. Plasmide, mit denen das die erfindungsgemäße Nukleinsäure aufweisende Genkonstrukt vorzugsweise in den Wirtsorganismus kloniert wird, sind weiter oben angegeben.
- Ein folgender Aspekt der Erfindung richtet sich auf Primer zur Herstellung der erfindungsgemäßen Gensequenzen mittels aller Arten von PCR. Mitumfasst sind die Sense- und Antisense-Primer codierend für die entsprechenden Aminosäuresequenzen, bzw. komplementären DNA-Sequenzen. Geeignete Primer können prinzipiell nach dem Fachmann bekannten Verfahren gewonnen werden. Das Auffinden der erfindungsgemäßen Primer erfolgt durch Vergleich mit bekannten DNA-Sequenzen oder durch Übersetzung der ins Auge gefaßten Aminosäuresequenzen in das bevorzugte Codon des betrachteten Organismus (z.B. für *Streptomyces*: Wright F. und Bibb M. J. (1992), *Codon usage in the G+C-rich Streptomyces genome*, *Gene* 113, 55-65). Gemeinsamkeiten in der Aminosäuresequenz von Proteinen von sogenannten Superfamilien sind hierfür ebenfalls von Nutzen (Firestine, S. M.; Nixon, A. E.; Benkovic, S. J. (1996), *Threading your way to protein function*, *Chem. Biol.* 3, 779-783). Weitere

Informationen diesbezüglich können gefunden werden in Gait, M. J. (1984), Oligonucleotide synthesis: a practical approach, IRL Press Ltd., Oxford; Innis, M. A.; Gelfound, D. H.; Sninsky, J. J. und White, T.J. (1990), PCR

5 Protocols: A guide to methods and applications, Academic Press Inc., San Diego.

Bevorzugte Primer sind die der Seq.ID.Nr. 11 und 12.

Für die Anwendung können die betrachteten Enzyme (Hydantoinrazemase, Hydantoinasen und/oder Carbamoylasen)

10 wie schon angedeutet in freier Form als homogen aufgereinigte Verbindungen oder als rekombinant (rec-) hergestelltes Enzym verwendet werden. Weiterhin können die Enzyme auch als Bestandteil eines intakten Gastorganismus eingesetzt werden oder in Verbindung mit der

15 aufgeschlossenen und beliebig hoch aufgereinigten Zellmasse des Wirtsorganismus.

Möglich ist ebenfalls die Verwendung der Enzyme in immobilisierter Form (Sharma B. P.; Bailey L. F. und

20 Messing R. A. (1982), Immobilisierte Biomaterialien - Techniken und Anwendungen, Angew. Chem. 94, 836-852). Vorteilhafterweise erfolgt die Immobilisierung durch Lyophilisation (Paradkar, V. M.; Dordick, J. S. (1994), Aqueous-Like Activity of α -Chymotrypsin Dissolved in Nearly Anhydrous Organic Solvents, J. Am. Chem. Soc. 116, 5009-5010; Mori, T.; Okahata, Y. (1997), A variety of lipi-coated glycoside hydrolases as effective glycosyl transfer catalysts in homogeneous organic solvents, Tetrahedron Lett. 38, 1971-1974; Otamiri, M.; Adlercreutz, P.; Matthiasson, B. (1992), Complex formation between

30 chymotrypsin and ethyl cellulose as a means to solubilize the enzyme in active form in toluene, Biocatalysis 6, 291-305). Ganz besonders bevorzugt ist die Lyophilisation in Gegenwart von oberflächenaktiven Substanzen, wie Aerosol OT oder Polyvinylpyrrolidon oder Polyethylenglycol (PEG) oder

35 Brij 52 (Diethylenglycol-mono-cetyl ether) (Kamiya, N.; Okazaki, S.-Y.; Goto, M. (1997), Surfactant-horseradish

peroxidase complex catalytically active in anhydrous benzene, Biotechnol. Tech. 11, 375-378).

Äußerst bevorzugt ist die Immobilisierung an Eupergit[®], insbesondere Eupergit C[®] und Eupergit 250L[®] (Röhm)

- 5 (Eupergit.RTM. C, a carrier for immobilization of enzymes of industrial potential. Katchalski-Katzir, E.; Kraemer, D. M. Journal of Molecular Catalysis B: Enzymatic (2000), 10(1-3), 157-176.)

- 10 Gleichfalls bevorzugt ist die Immobilisierung an Ni-NTA in Kombination mit dem His-Tag (Hexa-Histidin) ergänzten Polypeptid (Purification of proteins using polyhistidine affinity tags. Bornhorst, Joshua A.; Falke, Joseph J. Methods in Enzymology (2000), 326, 245-254). Die Verwendung als CLECs ist ebenfalls denkbar (St. Clair, N.; Wang, Y.-F.; Margolin, A. L. (2000), Cofactor-bound cross-linked enzyme crystals (CLEC) of alcohol dehydrogenase, Angew. Chem. Int. Ed. 39, 380-383).
- 15

- Durch diese Maßnahmen kann es gelingen aus Polypeptiden, welche durch organische Solventien instabil werden, solche zu generieren, die in Gemischen von wässrigen und organischen Lösungsmitteln bzw. ganz in Organik stabil sind und arbeiten können.
- 20

- Ganzzellkatalysatoren werden im Allgemeinen in Form freier oder immobilisierter Zellen eingesetzt. Hierzu wird die aktive Zellmasse in einer hydantoinhaltigen Lösung resuspendiert. Die Zellkonzentration beträgt dabei zwischen 1-100g/l. Die Konzentration des Hydantoins liegt zwischen 0,1 und 2 molar. Als Lösungsmittel wird bevorzugt H₂O verwendet, wobei jedoch auch Mischungen von organischen Lösungsmitteln und H₂O einsetzbar sind. Der pH-Wert wird entweder nicht geregelt oder mittels gängiger Puffer bzw. durch kontinuierliche pH-Statistierung zwischen pH6 und pH10 konstant gehalten. Die Reaktionstemperatur liegt typischerweise zwischen 20°C und 90°C. In Abhängigkeit der verwendeten Hydantoinase werden zweiwertige Metall-Ionen in Konzentrationen von 0,1-5mM hinzugesetzt. Bevorzugte
- 30
- 35

Metallionen sind dabei Mn^{2+} , Zn^{2+} oder Co^{2+} .

In Bezug auf den Einsatz der einzelnen Enzyme kann in äquivalenter Art und Weise verfahren werden.

Die durch den Einsatz der erfindungsgemäßen

- 5 Hydantoinrazemasen in wie z.B. oben beschriebener Weise hergestellten Produkte werden nach gängigen Verfahren aufgearbeitet. Vorteilhaft ist jedoch die Aufarbeitung durch Ionenaustauschchromatographie. Hierdurch wird das Produkt vom bei der Reaktion entstehenden Salzen befreit.
- 10 Das Eluat wird ggf. mit Aktivkohle geklärt und die entstandene enantiomerenangereicherte Aminosäure oder N-Carbamoyl-Aminosäure durch Einengung des Lösungsmittels ausgefällt und getrocknet.

- 15 Die Kopplung einer enzymatischen Razemisierung mit einer enantioselektiven Hydrolyse zum Screenen von Hydantoinrazemaseaktivitäten wurde bisher nicht zur Erzeugung verbesserter Hydantoinrazemasen angewendet. Für eine besonders erfolgreiche Anwendung des erfindungsgemäßen Verfahrens sollten mehrere Vorraussetzungen erfüllt sein:

- 20 1. Die chemische Razemisierungsgeschwindigkeit des im Screening verwendeten enantiomerenreinen Hydantoins muss sehr viel kleiner sein, als die Geschwindigkeit der enzymatisch katalysierten Reaktion.
- 25 2. Die enantioselektive enzymatische Hydrolyse mittels der Hydantoinase muss sehr viel schneller erfolgen als die enzymatische Razemisierung des Hydantoins.

- Für aliphatisch substituierte Hydantoine ist, durch deren langsame chemische Razemisierung bedingt, Punkt 1 gegeben. Punkt 2 kann durch eine gezielte Auswahl von geeigneten
- 30 Hydantoinasen (s. weiter vorne) erfüllt werden.

Mit den Aussagen des Standes der Technik wird die vorliegende Erfindung nicht nahegelegt, da diesem keinerlei

Hinweise auf die weiter oben genannten Voraussetzungen zu entnehmen sind.

Sämtliche der gezeigten Mutanten weisen an der Aminosäureposition 79 eine Mutation auf, was die Bedeutung dieser Position für die Enzymfunktion erstmalig aufzeigt. Interessant ist, dass die Aminosäuren, welche diese Position umgeben, für sämtliche bekannten Hydantoinrazemasen vollständig konserviert sind. Hieraus ergibt sich, dass für andere Hydantoinrazemasen welche das oben beschriebene Sequenzmotif enthalten und eine hohe Homologie (>40% Sequenzidentität) aufweisen durch ortsspezifische Mutagenese an Pos. 79 verbesserte Enzymvarianten erzeugt werden können, was bisher aus dem Stand der Technik nicht herleitbar war.

Unter optisch angereicherten (enantiomerenangereicherten, enantiomer angereicherten) Verbindungen wird im Rahmen der Erfindung das Vorliegen einer optischen Antipode im Gemisch mit der anderen in >50 mol-% verstanden.

Unter dem Begriff Nukleinsäuresequenzen werden alle Arten von einzelsträngiger oder doppelsträngiger DNA als auch RNA oder Gemische derselben subsumiert.

Die Verbesserung der Aktivität und/oder Selektivität und/oder Stabilität bedeutet erfindungsgemäß, dass die Polypeptide aktiver und/oder selektiver bzw. weniger selektiv oder unter den verwendeten Reaktionsbedingungen stabiler sind. Während die Aktivität und die Stabilität der Enzyme für die technische Anwendung naturgemäß möglichst hoch sein sollte, ist in Bezug auf die Selektivität dann von einer Verbesserung die Rede, wenn entweder die Substratselektivität abnimmt, die Enantioselektivität der Enzyme jedoch gesteigert ist.

Von den beanspruchten Polypeptiden und den Nukleinsäuresequenzen werden erfindungsgemäß auch solche Sequenzen umfaßt, die eine Homologie (exclusive der

natürlichen Degeneration) größer als 70% (in Bezug auf die Nukleinsäuresequenz) bzw. > 40% oder 80% (in Bezug auf die Polypeptide), bevorzugt größer als 90%, 91%, 92%, 93% oder 94%, mehr bevorzugt größer als 95% oder 96% und besonders
5 bevorzugt größer als 97%, 98% oder 99% zu einer dieser Sequenzen aufweisen, sofern die Wirkungsweise bzw. Zweck einer solchen Sequenz erhalten bleibt. Der Ausdruck "Homologie" (oder Identität) wie hierin verwendet, kann durch die Gleichung $H (\%) = [1 - V/X] \times 100$ definiert
10 werden, worin H Homologie bedeutet, X die Gesamtzahl an Nukleobasen/Aminosäuren der Vergleichssequenz ist und V die Anzahl an unterschiedlichen Nukleobasen/Aminosäuren der zu betrachtenden Sequenz bezogen auf die Vergleichssequenz ist. Auf jeden Fall sind mit dem Begriff Nukleinsäuresequenzen,
15 welche für Polypeptide codieren, alle Sequenzen umfaßt, die nach Maßgabe der Degeneration des genetischen Codes möglich erscheinen.

Der Ausdruck "unter stringenten Bedingungen" wird hierin wie bei Sambrook et al. (Sambrook, J.; Fritsch, E. F. und
20 Maniatis, T. (1989), Molecular cloning: a laboratory manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York) beschrieben, verstanden. Bevorzugt liegt eine stringente Hybridisierung gemäß der vorliegenden Erfindung vor, wenn nach Waschen für eine Stunde mit 1 x SSC (150 mM Natriumchlorid, 15 mM Natriumcitrat, pH 7.0) und 0,1 % SDS (Natriumdodecylsulfat) bei 50 °C, bevorzugt bei 55 °C, mehr bevorzugt bei 62 °C und am meisten bevorzugt bei 68 °C und mehr bevorzugt für 1 Stunde mit 0,2 x SSC und 0,1 % SDS bei
50 °C, bevorzugter bei 55 °C, mehr bevorzugt bei 62 °C und am
30 meisten bevorzugt bei 68 °C noch ein positives Hybridisierungssignal beobachtet wird.

Die in dieser Schrift genannten Literaturstellen gelten als von der Offenbarung mitumfaßt.

Der Organismus *Arthrobacter aureus* DSM3747 wurde durch
35 die Rütgerswerke Aktiengesellschaft am 28.05.86 bei der

Deutschen Sammlung für Mikroorganismen GmbH, Mascheroder
Weg 1b, 38124 Braunschweig hinterlegt.

Beispiele**Beispiel 1: Erzeugung Hydantoinrazemasemutanten - Zufallsmutagenese**

- 5 0,25ng des Vektors pOM21 (Plasmidkarte siehe Fig.1; Sequenz siehe Seq.ID.Nr.13) (PCT/US00/08159) wurde als Template in einem 100µl PCR Reaktionsmix bestehend aus PCR-Puffer (10 mM Tris, 1.5 mM MgCl₂, 50 mM KCl, pH 8.5), 200 µM dTTP, 200 µM dGTP, 200 µM dATP, 200 µM dCTP, 50 pmol des jeweiligen
- 10 Primers (siehe Seq.ID.Nr.11 und 12) und 2,5 U Taq-Polymerase (Roche) eingesetzt. Nach 30 Zyklen wurde das Amplifikat mittels Gelextraktion (QiaexII Gel-Extraktionskit) aufgereinigt und in den Vektor pOM21 mittels den Restriktionsenzymen NdeI und PstI subkloniert.
- 15 Das Ligationsprodukt wurde zur Transformation von hydantoinasepositiver Stämme verwendet (siehe Beispiel 2).

Beispiel 2: Herstellung von hydantoinasepositiven Stämmen und einer Mutantenbank

- 20 Chemisch kompetente *E.coli* JM109 (z.B. von Promega) wurden mit 10ng des Plasmids pDHYD (siehe Fig.2; siehe Seq.ID.Nr. 15) (**Herstellung?**) transformiert, welches das D-Hydantoinasegen aus *Arthrobacter crystallopoietes* DSM20117 unter Kontrolle eines Rhamnose-Promotors trägt. Die vollständige Sequenz des Plasmids ist in Seq.ID.Nr. 15
- 25 angegeben. Der so erzeugte hydantoinasepositive Stamm wurde wiederum chemisch kompetent gemacht und zur Herstellung der Mutantenbank mit dem Ligationsprodukt der Hydantoinrazemase-Zufallsmutagenese aus Beispiel 1 transformiert. Die auf Ampicillin- und Chloramphenicol-
- 30 haltigen Agarplatten ausgestrichenen Kolonien der Mutantenbank wurden anschliessend einem Screening unterworfen, welches in Beispiel 3 beschrieben wird.

Beispiel 3: Screening nach Hydantoinrazemasemutanten mit verbesserten Enzymeigenschaften

Einzelne Kolonien der Mutantenbank wurden in 96-Well-Platten überimpft, welche mit 100µl pro Well Rhamnose (2g/l) und ZnCl₂ (1mM) supplementiertem LB-Medium (5g/l Hefeextrakt, 10g/l Trypton, 10g/l NaCl) gefüllt waren. Die Platten wurden für 20 Stunden bei 30°C inkubiert.

Anschliessen wurden 100µl Screening-Substrat (100mM L-Ethylhydantoin, 50mg/l Cresol Rot, pH 8.5) zu jedem Well zugegeben und die Platten für 4 Stunden bei 20°C inkubiert. Wells mit verbesserten Hydantoinrazemasemutanten konnten durch eine intensivere Gelbfärbung im Vergleich zum Wildtyp direkt per Auge, oder unter Verwendung eines Spektralphotometers bei 580nm identifiziert werden.

Beispiel 4: Charakterisierung von Hydantoinrazemasemutanten mit verbesserten Enzymeigenschaften

Die im Screening identifizierten Razemasemutanten wurden anschliessend mittels HPLC-Analyse auf ihre Aktivität im Vergleich zum Wildtyp untersucht und die entsprechenden Mutationen mittels Sequenzierung bestimmt. Hierzu wurde von einzelnen Kolonien der unterschiedlichen Klone Plasmide isoliert (Qiagen Mini-Prep Kit) und sequenziert. Die selben Klone wurden zur Herstellung aktiver Biomasse verwendet.

Eine Übernachtskultur (OD₆₀₀=4) der jeweiligen Klone wurde hierzu 1:100 in 100ml Rhamnose (2g/l) und ZnCl₂ (1mM) supplementiertem LB-Medium (5g/l Hefeextrakt, 10g/l Trypton, 10g/l NaCl) verdünnt und 18 Stunden bei 30°C und 250UPM inkubiert. Die Biomasse wurde mittels Zentrifugation (10min, 10.000g) pelletiert und der Überstand verworfen. 2g aktive Biomasse wurde anschliessend in 50ml der Substratlösung (100mM L-Ethylhydantoin, pH 8.5) resuspendiert und bei 37°C inkubiert. Nach verschiedenen Zeiten wurden Proben genommen, die Biomasse durch Zentrifugation (5min, 13.000 UPM) abgetrennt und der

Überstand mittels HPLC auf die Konzentration der entstandenen N-Carbamoyl-aminobuttersäure analysiert.

Beispiel 5 Herstellung von L-Aminosäuren unter Verwendung verbesserter Hydantoinrazemasen

- 5 Ein mit pOM21-BB5 und pOM22 Fig. 3 (siehe Seq.ID.Nr.14) (PCT/US00/08159) transformierter Stamm von *E.coli* JM109 wurde bei 30°C in Ampicillin (100µg/l) und Chloramphenicol (50µg/l)-haltigem sowie mit 2g/l Rhamnose versetztem LB-Medium für 18 Stunden unter Schütteln (250 U/min)
- 10 inkubiert. Die Biomasse wurde durch Zentrifugation pelletiert und mit einem entsprechenden Volumen von 100mM DL-Ethlyhydantoinlösung, pH 8,5 und 1mM CoCl₂ so resuspendiert, dass sich eine Zellkonzentration von 30g/l ergibt. Diese Reaktionslösung wurde für 10 Stunden bei 37°C
- 15 inkubiert. Anschliessend wurden die Zellen durch Zentrifugation (30 min, 5000g) abgetrennt und der klare Überstand mittels HPLC auf die entstandene Aminosäure analysiert. Zur Aufarbeitung der enstandenen Aminosäure wurde das Volumen des Überstandes auf die Hälfte reduziert
- 20 und 1:2 mit Methanol versetzt. Die ausgefällte Aminosäure wurde anschliessend filtriert und getrocknet. Die Gesamtausbeute der Aminosäure betrug >60%.

Beispiel 6 Herstellung von D-Aminosäuren unter Verwendung verbesserter Hydantoinrazemasen

- 25 Ein mit pOM21-BB5 und pJAVIER16 Fig. 4 (siehe Seq.ID.Nr.16) (**Herstellung?**) transformierter Stamm von *E.coli* JM109 wurde bei 30°C in Ampicillin (100µg/l) und Chloramphenicol (50µg/l)-haltigem sowie mit 2g/l Rhamnose versetztem LB-Medium für 18 Stunden unter Schütteln (250 U/min)
- 30 inkubiert. Die Biomasse wurde durch Zentrifugation pelletiert und mit einem entsprechenden Volumen von 100mM DL-Ethlyhydantoinlösung, pH 8,5 und 1mM CoCl₂ so resuspendiert, dass sich eine Zellkonzentration von 30g/l ergibt. Diese Reaktionslösung wurde für 10 Stunden bei 37°C

inkubiert. Anschliessend wurden die Zellen durch
Zentrifugation (30 min, 5000g) abgetrennt und der klare
Überstand mittels HPLC auf die entstandene Aminosäure
analysiert. Zur Aufarbeitung der entstandenen Aminosäure
5 wurde das Volumen des Überstandes auf die Hälfte reduziert
und 1:2 mit Methanol versetzt. Die ausgefällte Aminosäure
wurde anschliessend filtriert und getrocknet. Die
Gesamtausbeute der Aminosäure betrug >60%.

Patentansprüche:

1. Screeningverfahren für Hydantoinrazemasen,
dadurch gekennzeichnet, dass man
 - a) eine enantioselektive Hydantoinase und
 - 5 b) die zu prüfende Hydantoinrazemase, welche eine
verglichen mit der Hydantoinase unter a) langsamere
Umsetzungsrate aufweist, auf
 - c) ein chirales Hydantoin einwirken
lässt, welches in zur Selektivität der Hydantoinase
10 entgegengesetzter enantiomerenangereicherten Form
eingesetzt wird, und
 - d) die resultierende N-Carbamoyl-Aminosäure oder die
freigesetzten Protonen zeitabhängig detektiert.
2. Verfahren nach Anspruch 1,
15 dadurch gekennzeichnet, dass man
ein aliphatisch substituiertes Hydantoin einsetzt.
3. Verfahren nach einem oder mehreren der vorhergehenden
Ansprüche,
dadurch gekennzeichnet, dass man
20 ein Hydantoinase aus *Arthrobacter crystallopoietes*
einsetzt.
4. Verfahren nach einem oder mehreren der vorhergehenden
Ansprüche,
dadurch gekennzeichnet, dass
25 das Verhältnis der Geschwindigkeitskonstanten der
Hydantoinase zur Hydantoinrazemase ($k_{\text{Hyd}}/k_{\text{Raz}}$) > 2 ist.
5. Verfahren zur Herstellung von verbesserten
Hydantoinrazemasen,
dadurch gekennzeichnet, dass man
 - 30 a) die Nukleinsäuresequenz codierend für die
Hydantoinrazemase einer Mutagenese unterwirft,
 - b) die aus a) erhältlichen Nukleinsäuresequenzen in
einen geeigneten Vektor kloniert und diesen in ein

geeignetes Expressionssystem transferiert und
c) die gebildeten Hydantoinrazemasen mit verbesserter
Aktivität und/oder Selektivität und/oder Stabilität
mittels eines Verfahrens nach einem oder mehreren der
Ansprüche 1 bis 4 detektiert und isoliert.

6. rec-Polypeptide oder diese codierende
Nukleinsäuresequenzen erhältlich nach Anspruch 5.
7. Verwendung der Polypeptide gemäß Anspruch 6 zur
Herstellung von enantiomerenangereicherten N-
Carbamoyl-Aminosäure oder Aminosäuren.
8. Verwendung der Nukleinsäuresequenzen gemäß 6 zur
Herstellung von Ganzzellkatalysatoren.
9. Hydantoinrazemase aufweisend in Position 79 einen
Aminosäureaustausch mit einer Aminosäure ausgewählt
aus der Gruppe bestehend aus A, R, N, D, C, Q, E, H,
I, L, K, M, F, P, S, T, Y oder V.
10. Hydantoinrazemasen aufweisend die Konsensussequenz
FX₁DX₂GL (Seq. 1), wobei X₂ P oder T darstellt und X₁
in der Position 79 eine Aminosäure ausgewählt aus der
Gruppe A, R, N, D, C, Q, E, H, I, L, K, M, F, P, S, T,
Y oder V darstellt darstellt.
11. Isolierte Nukleinsäuresequenz codierend für eine
Hydantoinrazemase ausgewählt aus der Gruppe:
 - a) einer Nukleinsäuresequenz codierend für eine
Hydantoinrazemase gemäß Anspruch 9 und/oder 10,
 - b) einer Nukleinsäuresequenz, die unter stringenten
Bedingungen mit der Nukleinsäuresequenz codierend
für eine Hydantoinrazemase gemäß Anspruch 9
und/oder 10 oder der dazu komplementären Sequenz
hybridisiert,
 - c) einer Nukleinsäuresequenz gemäß den Seq.ID.Nr. 3,
5, 7 oder 9 oder solchen mit einer Homologie von
> 80% zu diesen,

d) einer Nukleinsäuresequenz aufweisend 15
aufeinanderfolgende Nukleotide der Sequenzen
Seq.ID.Nr. 3, 5, 7 oder 9..

- 5
12. Ganzzellkatalysator aufweisend ein kloniertes Gen für
eine Hydantoinrazemase gemäß den Ansprüchen 9 und/oder
10.
13. Plasmide, Vektoren oder Mikroorganismen aufweisend
eine Nukleinsäuresequenz gemäß Anspruch 9 und/oder 10.
14. Primer zur Herstellung der Nukleinsäuresequenzen nach
Anspruch 9 und/oder 10 mittels PCR.
- 0

Zusammenfassung:

Die vorliegende Erfindung bezieht sich auf ein Screeningverfahren für Hydantoinrazemasen und neue Hydantoinrazemasen, die sie codierenden

5 Nukleinsäuresequenzen und ein Verfahren zur Mutagenese.

Hydantoinrazemasen sind im Zusammenhang mit der Erzeugung von enantiomerenangereicherten Aminosäuren aus racemischen Hydantoinen von Interesse.

Abb. 1:

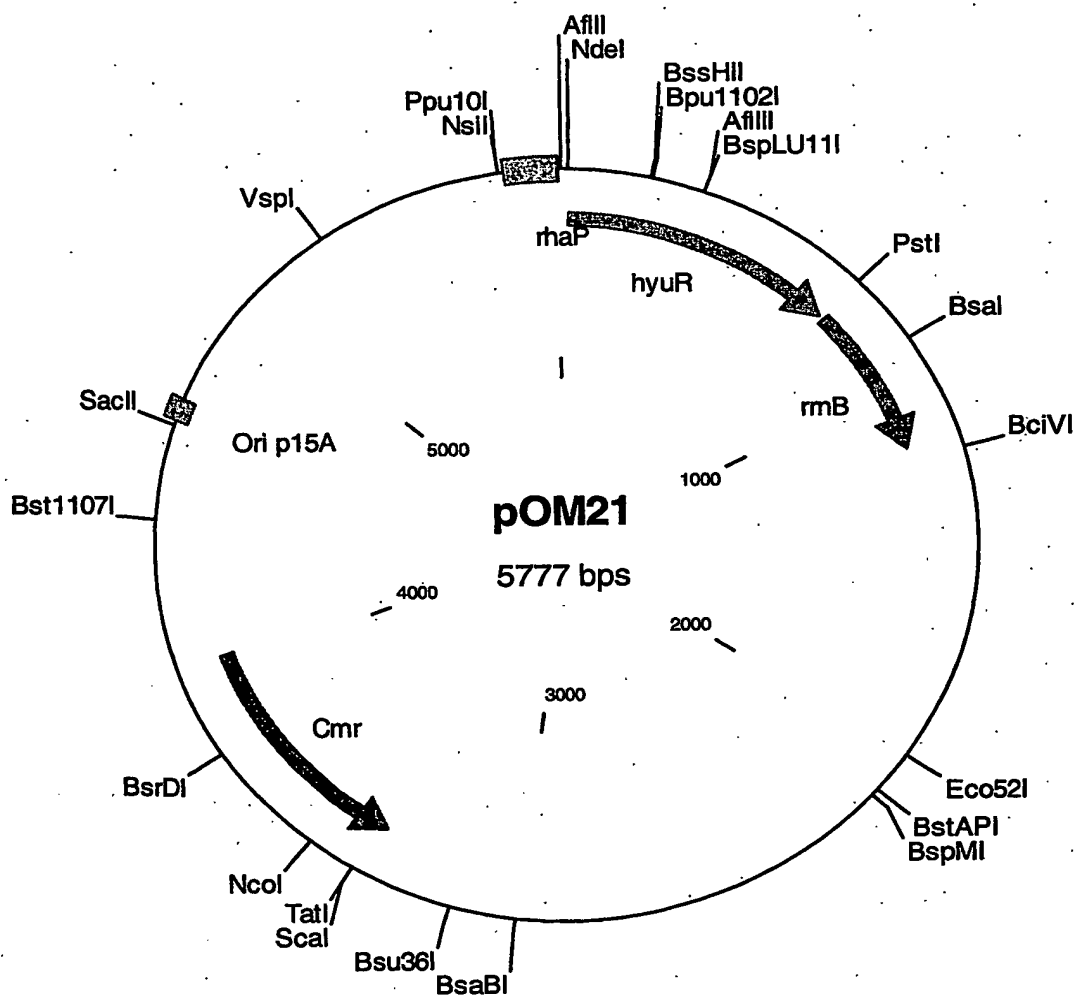


Abb. 2:

5

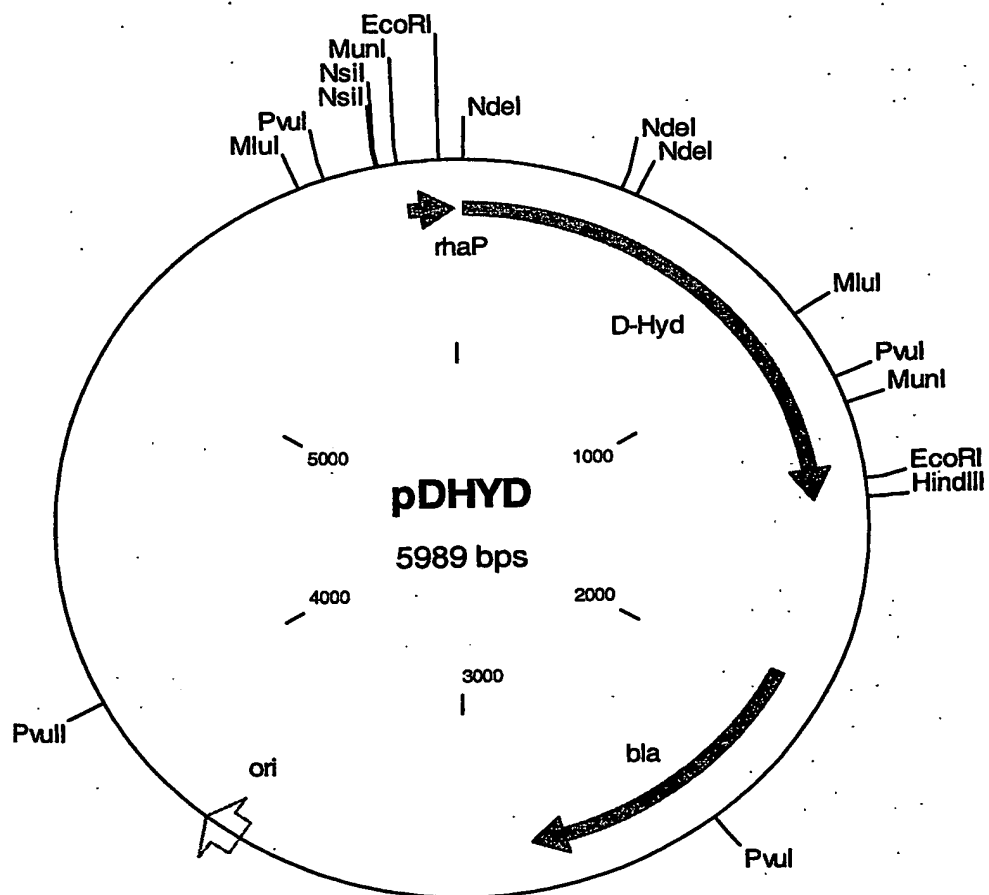


Fig: 3

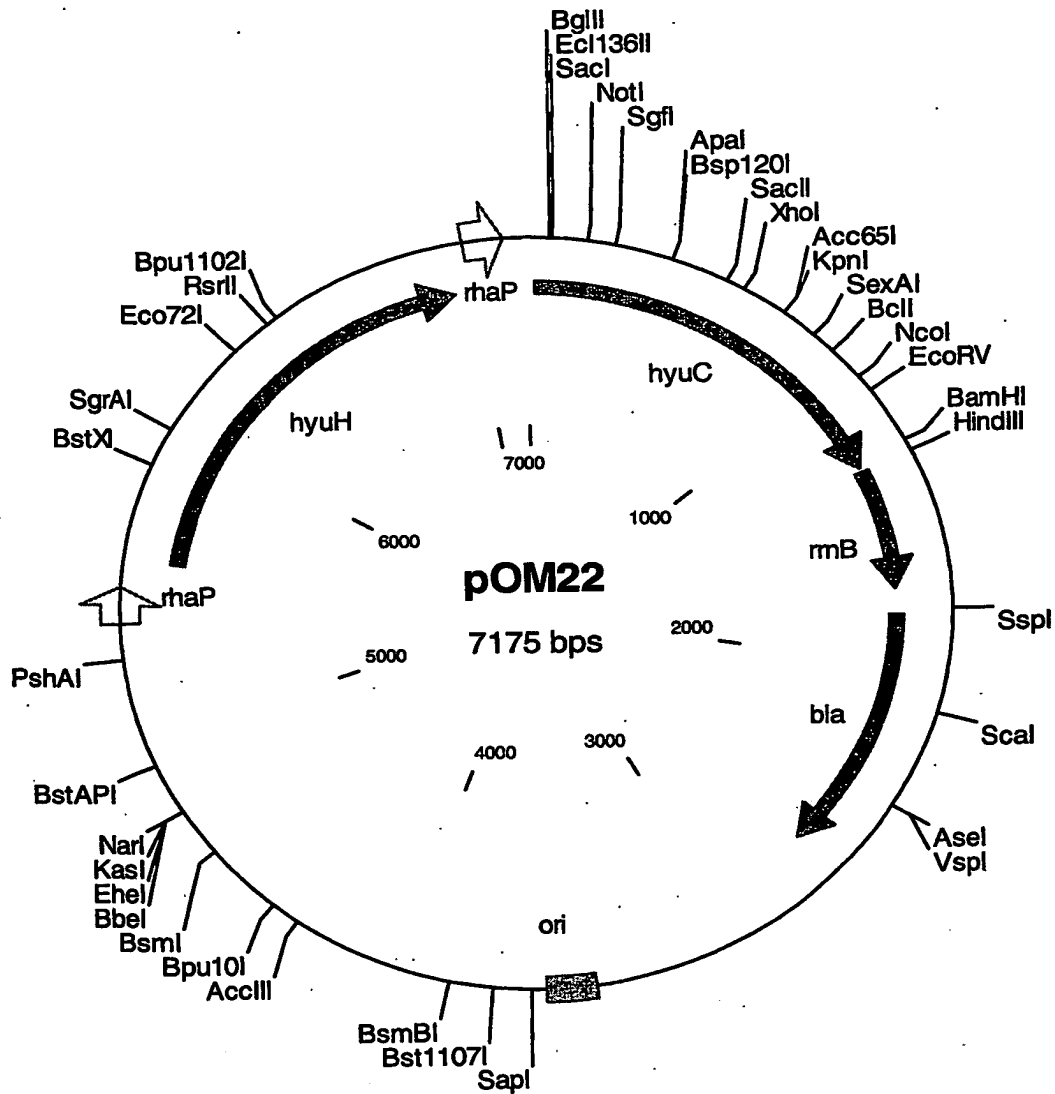
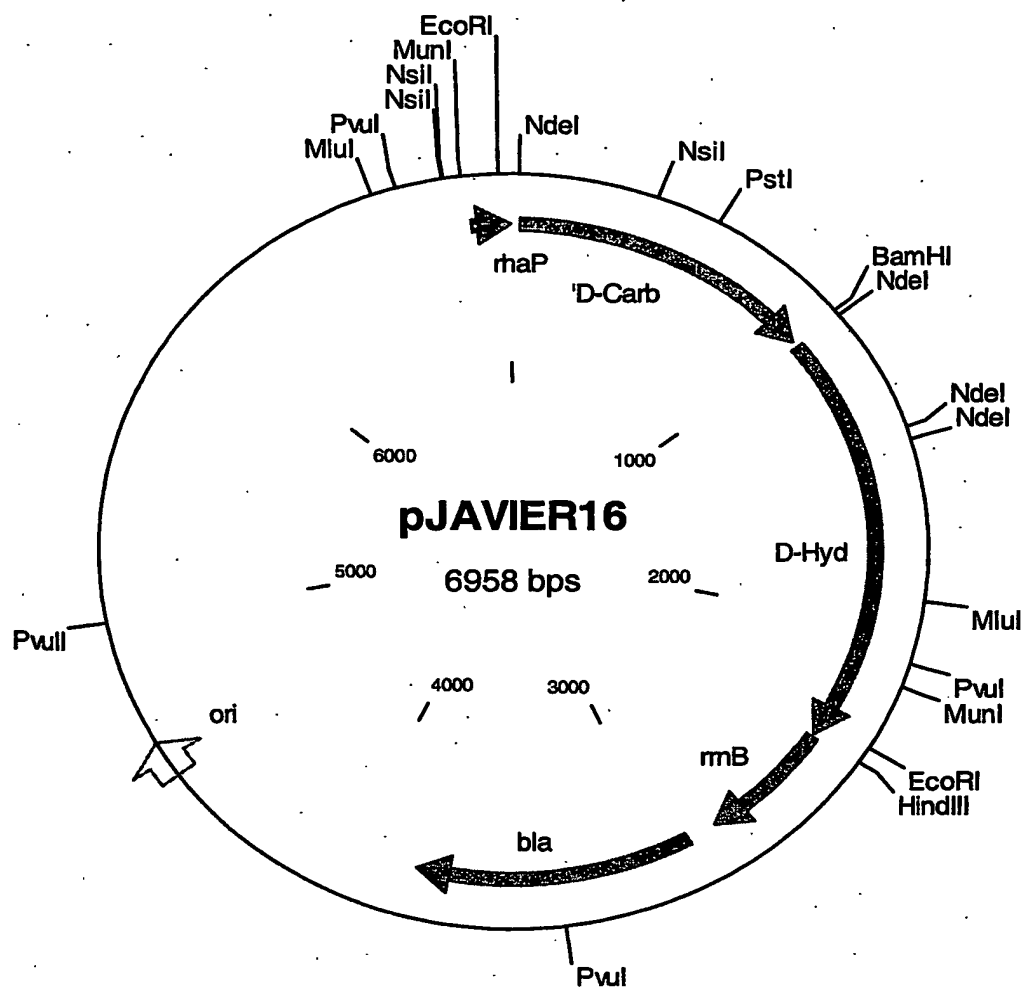


Fig. 4



SEQUENZPROTOKOLL

<110> Degussa AG

5 <120> Screeningverfahren für Hydantoinrazemasen

<130> 030115 AM / IP

<140>

10 <141>

<160> 16

<170> PatentIn Ver. 2.1

15

<210> 1

<211> 6

<212> PRT

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:
Konsensussequenz

25

<400> 1

Phe Xaa Asp Xaa Gly Leu
1 5

30

<210> 2

<211> 236

<212> PRT

<213> Arthrobacter crystallopoietes

35

<400> 2

Met Arg Ile Leu Val Ile Asn Pro Asn Ser Ser Ser Ala Leu Thr Glu
1 5 10 15Ser Val Ala Asp Ala Ala Gln Gln Val Val Ala Thr Gly Thr Ile Ile
20 25 30Ser Ala Ile Asn Pro Ser Arg Gly Pro Ala Val Ile Glu Gly Ser Phe
35 40 4545 Asp Glu Ala Leu Ala Thr Phe His Leu Ile Glu Glu Val Glu Arg Ala
50 55 60Glu Arg Glu Asn Pro Pro Asp Ala Tyr Val Ile Ala Cys Phe Gly Asp
65 70 75 80

50

Pro Gly Leu Asp Ala Val Lys Glu Leu Thr Asp Arg Pro Val Val Gly
85 90 95

55

Val Ala Glu Ala Ala Ile His Met Ser Ser Phe Val Ala Ala Thr Phe
100 105 110Ser Ile Val Ser Ile Leu Pro Arg Val Arg Lys His Leu His Glu Leu
115 120 125

Val Arg Gln Ala Gly Ala Thr Asn Arg Leu Ala Ser Ile Lys Leu Pro
 130 135 140

5 Asn Leu Gly Val Met Ala Phe His Glu Asp Glu His Ala Ala Leu Glu
 145 150 155 160

Thr Leu Lys Gln Ala Ala Lys Glu Ala Val Gln Glu Asp Gly Ala Glu
 165 170 175

10 Ser Ile Val Leu Gly Cys Ala Gly Met Val Gly Phe Ala Arg Gln Leu
 180 185 190

Ser Asp Glu Leu Gly Val Pro Val Ile Asp Pro Val Glu Ala Ala Cys
 195 200 205

15 Arg Val Ala Glu Ser Leu Val Ala Leu Gly Tyr Gln Thr Ser Lys Ala
 210 215 220

20 Asn Ser Tyr Gln Lys Pro Thr Glu Lys Gln Tyr Leu
 225 230 235

<210> 3
 <211> 711
 25 <212> DNA
 <213> Künstliche Sequenz

<220>
 <223> Beschreibung der künstlichen Sequenz:1BG7
 30

<220>
 <221> CDS
 <222> (1)..(711)

35 <400> 3
 atg aga atc ctc gtg atc aac ccc aac agt tcc agc gcc ctt act gaa 48
 Met Arg Ile Leu Val Ile Asn Pro Asn Ser Ser Ser Ala Leu Thr Glu
 1 5 10 15

tcg gtt gcg gac gca gca caa caa gtt gtc gcg acc ggc acc ata att 96
 Ser Val Ala Asp Ala Ala Gln Gln Val Val Ala Thr Gly Thr Ile Ile
 20 25 30

tct gcc atc aac ccc tcc aga gga ccc gcc gtc att gaa ggc agc ttt 144
 45 Ser Ala Ile Asn Pro Ser Arg Gly Pro Ala Val Ile Glu Gly Ser Phe
 35 40 45

gac gaa gca ctg gcc acg ttc cat ctc att gaa gag gtg gag cgc gct 192
 50 Asp Glu Ala Leu Ala Thr Phe His Leu Ile Glu Glu Val Glu Arg Ala
 50 55 60

gag cgg gaa aac ccg ccc gac gcc tac gtc atc gca tgt ttc agg gat 240
 Glu Arg Glu Asn Pro Pro Asp Ala Tyr Val Ile Ala Cys Phe Arg Asp
 65 70 75 80

55 ccg gga ctt gac gcg gtc aag gag ctg act gac agg cca gtg gta gga 288
 Pro Gly Leu Asp Ala Val Lys Glu Leu Thr Asp Arg Pro Val Val Gly
 85 90 95

gtt gcc gaa gct gca atc cac atg tct tca ttc gtc gcg gcc acc ttc 336
 Val Ala Glu Ala Ala Ile His Met Ser Ser Phe Val Ala Ala Thr Phe
 100 105 110

5 tcc att gtc agc atc ctc ccg agg gtc agg aaa cat ctg cac gaa ctg 384
 Ser Ile Val Ser Ile Leu Pro Arg Val Arg Lys His Leu His Glu Leu
 115 120 125

10 gta cgg caa gcg ggg gcg acg aat cgc ctc gcc tcc atc aag ctc cca 432
 Val Arg Gln Ala Gly Ala Thr Asn Arg Leu Ala Ser Ile Lys Leu Pro
 130 135 140

15 aat ctg ggg gtg atg gcc ttc cat gag gac gaa cat gcc gca ctg gag 480
 Asn Leu Gly Val Met Ala Phe His Glu Asp Glu His Ala Ala Leu Glu
 145 150 155 160

0 acg ctc aaa caa gcc gcc aag gag gcg gtc cag gag gac ggc gcc gag 528
 Thr Leu Lys Gln Ala Ala Lys Glu Ala Val Gln Glu Asp Gly Ala Glu
 165 170 175

tcg ata gtg ctc gga tgc gcc ggc atg gtg ggg ttt gcg cgt caa ctg 576
 Ser Ile Val Leu Gly Cys Ala Gly Met Val Gly Phe Ala Arg Gln Leu
 180 185 190

25 agc gac gaa ctc ggc gtc cct gtc atc gac ccc gtc gag gca gct tgc 624
 Ser Asp Glu Leu Gly Val Pro Val Ile Asp Pro Val Glu Ala Ala Cys
 195 200 205

30 cgc gtg gcc gag agt ttg gtc gct ctg ggc tac cag acc agc aaa gcg 672
 Arg Val Ala Glu Ser Leu Val Ala Leu Gly Tyr Gln Thr Ser Lys Ala
 210 215 220

35 aac tcg tat caa aaa ccg aca gag aag cag tac ctc tag 711
 Asn Ser Tyr Gln Lys Pro Thr Glu Lys Gln Tyr Leu
 225 230 235

<210> 4
 <211> 237
 <212> PRT
 <213> Künstliche Sequenz
 <223> Beschreibung der künstlichen Sequenz:1BG7

45 <400> 4
 Met Arg Ile Leu Val Ile Asn Pro Asn Ser Ser Ser Ala Leu Thr Glu
 1 5 10 15

Ser Val Ala Asp Ala Ala Gln Gln Val Val Ala Thr Gly Thr Ile Ile
 20 25 30

50 Ser Ala Ile Asn Pro Ser Arg Gly Pro Ala Val Ile Glu Gly Ser Phe
 35 40 45

55 Asp Glu Ala Leu Ala Thr Phe His Leu Ile Glu Glu Val Glu Arg Ala
 50 55 60

Glu Arg Glu Asn Pro Pro Asp Ala Tyr Val Ile Ala Cys Phe Arg Asp
 65 70 75 80

Pro Gly Leu Asp Ala Val Lys Glu Leu Thr Asp Arg Pro Val Val Gly
85 90 95

5 Val Ala Glu Ala Ala Ile His Met Ser Ser Phe Val Ala Ala Thr Phe
100 105 110

Ser Ile Val Ser Ile Leu Pro Arg Val Arg Lys His Leu His Glu Leu
115 120 125

10 Val Arg Gln Ala Gly Ala Thr Asn Arg Leu Ala Ser Ile Lys Leu Pro
130 135 140

Asn Leu Gly Val Met Ala Phe His Glu Asp Glu His Ala Ala Leu Glu
145 150 155 160

15 Thr Leu Lys Gln Ala Ala Lys Glu Ala Val Gln Glu Asp Gly Ala Glu
165 170 175

0 Ser Ile Val Leu Gly Cys Ala Gly Met Val Gly Phe Ala Arg Gln Leu
180 185 190

Ser Asp Glu Leu Gly Val Pro Val Ile Asp Pro Val Glu Ala Ala Cys
195 200 205

25 Arg Val Ala Glu Ser Leu Val Ala Leu Gly Tyr Gln Thr Ser Lys Ala
210 215 220

Asn Ser Tyr Gln Lys Pro Thr Glu Lys Gln Tyr Leu
225 230 235

30

<210> 5
<211> 711
<212> DNA
35 <213> Künstliche Sequenz

<220>
<223> Beschreibung der künstlichen Sequenz:3CH11

<220>
<221> CDS
<222> (1)..(711)

<400> 5

45 atg aga atc ctc gtg atc aac ccc aac agt tcc agc gcc ctt act gaa 48
Met Arg Ile Leu Val Ile Asn Pro Asn Ser Ser Ser Ala Leu Thr Glu
1 5 10 15

50 tcg gtt gcg gac gca gca caa caa gtt gtc gcg acc ggc acc ata att 96
Ser Val Ala Asp Ala Ala Gln Gln Val Val Ala Thr Gly Thr Ile Ile
20 25 30

tct gcc atc aac ccc tcc aga gga ccc gcc gtc att gaa ggc agc ttt 144
Ser Ala Ile Asn Pro Ser Arg Gly Pro Ala Val Ile Glu Gly Ser Phe
35 40 45

55 gac gaa gca ctg gcc acg ttc cat ctc att gaa gag gtg gag cgc gct 192
Asp Glu Ala Leu Ala Thr Phe His Leu Ile Glu Glu Val Glu Arg Ala
50 55 60

gag cgg gaa aac ccg ccc gac gcc tac gtc atc gca tgt ttc gag gat 240
 Glu Arg Glu Asn Pro Pro Asp Ala Tyr Val Ile Ala Cys Phe Glu Asp
 65 70 75 80

5 ccg gga ctt gac gcg gtc aag gag ctg act gac agg cca gtg gta gga 288
 Pro Gly Leu Asp Ala Val Lys Glu Leu Thr Asp Arg Pro Val Val Gly
 85 90 95

10 gtt gcc gaa gct gca atc cac atg tct tca ttc gtc gcg gcc acc ttc 336
 Val Ala Glu Ala Ala Ile His Met Ser Ser Phe Val Ala Ala Thr Phe
 100 105 110

15 tcc att gtc agc atc ctc ccg agg gtc agg aaa cat ctg cac gaa ctg 384
 Ser Ile Val Ser Ile Leu Pro Arg Val Arg Lys His Leu His Glu Leu
 115 120 125

20 gta cgg caa gcg ggg gcg acg aat cgc ctc gcc tcc atc aag ctc cca 432
 Val Arg Gln Ala Gly Ala Thr Asn Arg Leu Ala Ser Ile Lys Leu Pro
 130 135 140

25 aat ctg ggg gtg atg gcc ttc cat gag gac gaa cat gcc gca ctg gag 480
 Asn Leu Gly Val Met Ala Phe His Glu Asp Glu His Ala Ala Leu Glu
 145 150 155 160

acg ctc aaa caa gcc gcc aag gag gcg gtc cag gag gac ggc gcc gag 528
 Thr Leu Lys Gln Ala Ala Lys Glu Ala Val Gln Glu Asp Gly Ala Glu
 165 170 175

30 tcg ata gtg ctc gga tgc gcc ggc atg gtg ggg ttt gcg cgt caa ctg 576
 Ser Ile Val Leu Gly Cys Ala Gly Met Val Gly Phe Ala Arg Gln Leu
 180 185 190

35 agc gac gaa ctc ggc gtc cct gtc atc gac ccc gtc gag gca gct tgc 624
 Ser Asp Glu Leu Gly Val Pro Val Ile Asp Pro Val Glu Ala Ala Cys
 195 200 205

40 cgc gtg gcc gag agt ttg gtc gct ctg ggc tac cag acc agc aaa gcg 672
 Arg Val Ala Glu Ser Leu Val Ala Leu Gly Tyr Gln Thr Ser Lys Ala
 210 215 220

45 aac tcg tat caa aaa ccg aca gag aag cag tac ctc tag 711
 Asn Ser Tyr Gln Lys Pro Thr Glu Lys Gln Tyr Leu
 225 230 235

50 <210> 6
 <211> 237
 <212> PRT
 <213> Künstliche Sequenz
 <223> Beschreibung der künstlichen Sequenz:3CH11

55 <400> 6
 Met Arg Ile Leu Val Ile Asn Pro Asn Ser Ser Ser Ala Leu Thr Glu
 1 5 10 15
 Ser Val Ala Asp Ala Ala Gln Gln Val Val Ala Thr Gly Thr Ile Ile
 20 25 30

Ser Ala Ile Asn Pro Ser Arg Gly Pro Ala Val Ile Glu Gly Ser Phe
 35 40 45

5 Asp Glu Ala Leu Ala Thr Phe His Leu Ile Glu Glu Val Glu Arg Ala
 50 55 60

Glu Arg Glu Asn Pro Pro Asp Ala Tyr Val Ile Ala Cys Phe Glu Asp
 65 70 75 80

10 Pro Gly Leu Asp Ala Val Lys Glu Leu Thr Asp Arg Pro Val Val Gly
 85 90 95

Val Ala Glu Ala Ala Ile His Met Ser Ser Phe Val Ala Ala Thr Phe
 100 105 110

15 Ser Ile Val Ser Ile Leu Pro Arg Val Arg Lys His Leu His Glu Leu
 115 120 125

20 Val Arg Gln Ala Gly Ala Thr Asn Arg Leu Ala Ser Ile Lys Leu Pro
 130 135 140

Asn Leu Gly Val Met Ala Phe His Glu Asp Glu His Ala Ala Leu Glu
 145 150 155 160

25 Thr Leu Lys Gln Ala Ala Lys Glu Ala Val Gln Glu Asp Gly Ala Glu
 165 170 175

Ser Ile Val Leu Gly Cys Ala Gly Met Val Gly Phe Ala Arg Gln Leu
 180 185 190

30 Ser Asp Glu Leu Gly Val Pro Val Ile Asp Pro Val Glu Ala Ala Cys
 195 200 205

35 Arg Val Ala Glu Ser Leu Val Ala Leu Gly Tyr Gln Thr Ser Lys Ala
 210 215 220

Asn Ser Tyr Gln Lys Pro Thr Glu Lys Gln Tyr Leu
 225 230 235

45 <210> 7
 <211> 711
 <212> DNA
 <213> Künstliche Sequenz

<220>
 <223> Beschreibung der künstlichen Sequenz:AE3

50 <220>
 <221> CDS
 <222> (1) .. (711)

55 <400> 7
 atg aga atc ctc gtg atc aac ccc aac agt tcc agc gcc ctt act gaa 48
 Met Arg Ile Leu Val Ile Asn Pro Asn Ser Ser Ser Ala Leu Thr Glu
 1 5 10 15

tcg gtt gcg gac gca gca caa caa gtt gtc gcg acc ggc acc ata att 96
 Ser Val Ala Asp Ala Ala Gln Gln Val Val Ala Thr Gly Thr Ile Ile

	20	25	30	
5	tct gcc atc aac ccc tcc aga gga ccc gcc gtc att gaa ggc agc ttt Ser Ala Ile Asn Pro Ser Arg Gly Pro Ala Val Ile Glu Gly Ser Phe	40	45	144
10	gac gaa gca ctg gcc acg ttc cat ctc att gaa gag gtg gag cgc gct Asp Glu Ala Leu Ala Thr Phe His Leu Ile Glu Glu Val Glu Arg Ala	55	60	192
15	gag cgg gaa aac ccg ccc gac gcc tac gtc atc gca tgt ttc cag gat Glu Arg Glu Asn Pro Pro Asp Ala Tyr Val Ile Ala Cys Phe Gln Asp	70	75	240
20	ccg gga ctt gac gcg gtc aag gag ctg act gac agg cca gtg gta gga Pro Gly Leu Asp Ala Val Lys Glu Leu Thr Asp Arg Pro Val Val Gly	85	90	288
25	gtt gcc gaa gct gca atc cac atg tct tca ttc gtc gcg gcc acc ttc Val Ala Glu Ala Ala Ile His Met Ser Ser Phe Val Ala Ala Thr Phe	100	105	336
30	tcc att gtc agc atc ctc ccg agg gtc agg aaa cat ctg cac gaa ctg Ser Ile Val Ser Ile Leu Pro Arg Val Arg Lys His Leu His Glu Leu	115	120	384
35	gta cgg caa gcg ggg gcg acg aat cgc ctc gcc tcc atc aag ctc cca Val Arg Gln Ala Gly Ala Thr Asn Arg Leu Ala Ser Ile Lys Leu Pro	130	135	432
40	aat ctg ggg gtg atg gcc ttc cat gag gac gaa cat gcc gca ctg gag Asn Leu Gly Val Met Ala Phe His Glu Asp Glu His Ala Ala Leu Glu	145	150	480
45	acg ctc aaa caa gcc gcc aag gag gcg gtc cag gag gac ggc gcc gag Thr Leu Lys Gln Ala Ala Lys Glu Ala Val Gln Glu Asp Gly Ala Glu	165	170	528
50	tcg ata gtg ctc gga tgc gcc ggc atg gtg ggg ttt gcg cgt caa ctg Ser Ile Val Leu Gly Cys Ala Gly Met Val Gly Phe Ala Arg Gln Leu	180	185	576
55	agc gac gaa ctc ggc gtc cct gtc atc gac ccc gtc gag gca gct tgc Ser Asp Glu Leu Gly Val Pro Val Ile Asp Pro Val Glu Ala Ala Cys	195	200	624
60	cgc gtg gcc gag agt ttg gtc gct ctg ggc tac cag acc agc aaa gcg Arg Val Ala Glu Ser Leu Val Ala Leu Gly Tyr Gln Thr Ser Lys Ala	210	215	672
65	aac tcg tat caa aaa ccg aca gag aag cag tac ctc tag Asn Ser Tyr Gln Lys Pro Thr Glu Lys Gln Tyr Leu	225	230	711
70				
75				
80				
85				
90				
95				
100				
105				
110				
115				
120				
125				
130				
135				
140				
145				
150				
155				
160				
165				
170				
175				
180				
185				
190				
195				
200				
205				
210				
215				
220				
225				
230				
235				
240				
245				
250				
255				
260				
265				
270				
275				
280				
285				
290				
295				
300				
305				
310				
315				
320				
325				
330				
335				
340				
345				
350				
355				
360				
365				
370				
375				
380				
385				
390				
395				
400				
405				
410				
415				
420				
425				
430				
435				
440				
445				
450				
455				
460				
465				
470				
475				
480				
485				
490				
495				
500				
505				
510				
515				
520				
525				
530				
535				
540				
545				
550				
555				
560				
565				
570				
575				
580				
585				
590				
595				
600				
605				
610				
615				
620				
625				
630				
635				
640				
645				
650				
655				
660				
665				
670				
675				
680				
685				
690				
695				
700				
705				
710				
715				
720				
725				
730				
735				
740				
745				
750				
755				
760				
765				
770				
775				
780				
785				
790				
795				
800				
805				
810				
815				
820				
825				
830				
835				
840				
845				
850				
855				
860				
865				
870				
875				
880				
885				
890				
895				
900				
905				
910				
915				
920				
925				
930				
935				
940				
945				
950				
955				
960				
965				
970				
975				
980				
985				
990				
995				

<210> 8
 <211> 237
 <212> PRT
 <213> Künstliche Sequenz

<223> Beschreibung der künstlichen Sequenz:AE3

<400> 8

5 Met Arg Ile Leu Val Ile Asn Pro Asn Ser Ser Ser Ala Leu Thr Glu
1 5 10 15

Ser Val Ala Asp Ala Ala Gln Gln Val Val Ala Thr Gly Thr Ile Ile
20 25 30

10 Ser Ala Ile Asn Pro Ser Arg Gly Pro Ala Val Ile Glu Gly Ser Phe
35 40 45

Asp Glu Ala Leu Ala Thr Phe His Leu Ile Glu Glu Val Glu Arg Ala
50 55 60

15 Glu Arg Glu Asn Pro Pro Asp Ala Tyr Val Ile Ala Cys Phe Gln Asp
65 70 75 80

0 Pro Gly Leu Asp Ala Val Lys Glu Leu Thr Asp Arg Pro Val Val Gly
85 90 95

Val Ala Glu Ala Ala Ile His Met Ser Ser Phe Val Ala Ala Thr Phe
100 105 110

25 Ser Ile Val Ser Ile Leu Pro Arg Val Arg Lys His Leu His Glu Leu
115 120 125

Val Arg Gln Ala Gly Ala Thr Asn Arg Leu Ala Ser Ile Lys Leu Pro
130 135 140

30 Asn Leu Gly Val Met Ala Phe His Glu Asp Glu His Ala Ala Leu Glu
145 150 155 160

35 Thr Leu Lys Gln Ala Ala Lys Glu Ala Val Gln Glu Asp Gly Ala Glu
165 170 175

Ser Ile Val Leu Gly Cys Ala Gly Met Val Gly Phe Ala Arg Gln Leu
180 185 190

Ser Asp Glu Leu Gly Val Pro Val Ile Asp Pro Val Glu Ala Ala Cys
195 200 205

Arg Val Ala Glu Ser Leu Val Ala Leu Gly Tyr Gln Thr Ser Lys Ala
210 215 220

45 Asn Ser Tyr Gln Lys Pro Thr Glu Lys Gln Tyr Leu
225 230 235

50 <210> 9

<211> 711

<212> DNA

<213> Künstliche Sequenz

55 <220>

<223> Beschreibung der künstlichen Sequenz:BB5

<220>

<221> CDS

<222> (1)..(711)

<400> 9

5	atg aga atc ctc gtg atc aac ccc aac agt tcc agc gcc ctt act gaa	48
	Met Arg Ile Leu Val Ile Asn Pro Asn Ser Ser Ser Ala Leu Thr Glu	
	1 5 10 15	
10	tcg gtt gcg gac gca gca caa caa gtt gtc gcg acc ggc acc ata att	96
	Ser Val Ala Asp Ala Ala Gln Gln Val Val Ala Thr Gly Thr Ile Ile	
	20 25 30	
15	tct gcc atc aac ccc tcc aga gga ccc gcc gtc att gaa ggc agc ttt	144
	Ser Ala Ile Asn Pro Ser Arg Gly Pro Ala Val Ile Glu Gly Ser Phe	
	35 40 45	
20	gac gaa gca ctg gcc acg ttc cat ctc att gaa gag gtg gag cgc gct	192
	Asp Glu Ala Leu Ala Thr Phe His Leu Ile Glu Glu Val Glu Arg Ala	
	50 55 60	
25	gag cgg gaa aac ccg ccc gac gcc tac gtc atc gca tgt ttc ttg gat	240
	Glu Arg Glu Asn Pro Pro Asp Ala Tyr Val Ile Ala Cys Phe Leu Asp	
	65 70 75 80	
30	ccg gga ctt gac gcg gtc aag gag ctg act gac agg cca gtg gta gga	288
	Pro Gly Leu Asp Ala Val Lys Glu Leu Thr Asp Arg Pro Val Val Gly	
	85 90 95	
35	gtt gcc gaa gct gca atc cac atg tct tca ttc gtc gcg gcc acc ttc	336
	Val Ala Glu Ala Ala Ile His Met Ser Ser Phe Val Ala Ala Thr Phe	
	100 105 110	
40	tcc att gtc agc atc ctc ccg agg gtc agg aaa cat ctg cac gaa ctg	384
	Ser Ile Val Ser Ile Leu Pro Arg Val Arg Lys His Leu His Glu Leu	
	115 120 125	
45	gta cgg caa gcg ggg gcg acg aat cgc ctc gcc tcc atc aag ctc cca	432
	Val Arg Gln Ala Gly Ala Thr Asn Arg Leu Ala Ser Ile Lys Leu Pro	
	130 135 140	
50	aat ctg ggg gtg atg gcc ttc cat gag gac gaa cat gcc gca ctg gag	480
	Asn Leu Gly Val Met Ala Phe His Glu Asp Glu His Ala Ala Leu Glu	
	145 150 155 160	
55	acg ctc aaa caa gcc gcc aag gag gcg gtc cag gag gac ggc gcc gag	528
	Thr Leu Lys Gln Ala Ala Lys Glu Ala Val Gln Glu Asp Gly Ala Glu	
	165 170 175	
60	tcg ata gtg ctc gga tgc gcc ggc atg gtg ggg ttt gcg cgt caa ctg	576
	Ser Ile Val Leu Gly Cys Ala Gly Met Val Gly Phe Ala Arg Gln Leu	
	180 185 190	
65	agc gac gaa ctc ggc gtc cct gtc atc gac ccc gtc gag gca gct tgc	624
	Ser Asp Glu Leu Gly Val Pro Val Ile Asp Pro Val Glu Ala Ala Cys	
	195 200 205	
70	cgc gtg gcc gag agt ttg gtc gct ctg ggc tac cag acc agc aaa gcg	672
	Arg Val Ala Glu Ser Leu Val Ala Leu Gly Tyr Gln Thr Ser Lys Ala	
	210 215 220	

aac tcg tat caa aaa ccg aca gag aag cag tac ctc tag
 Asn Ser Tyr Gln Lys Pro Thr Glu Lys Gln Tyr Leu
 225 230 235

711

5

<210> 10

<211> 237

<212> PRT

<213> Künstliche Sequenz

10 <223> Beschreibung der künstlichen Sequenz:BB5

<400> 10

Met Arg Ile Leu Val Ile Asn Pro Asn Ser Ser Ser Ala Leu Thr Glu
 1 5 10 15

15

Ser Val Ala Asp Ala Ala Gln Gln Val Val Ala Thr Gly Thr Ile Ile
 20 25 30

Ser Ala Ile Asn Pro Ser Arg Gly Pro Ala Val Ile Glu Gly Ser Phe
 35 40 45

Asp Glu Ala Leu Ala Thr Phe His Leu Ile Glu Glu Val Glu Arg Ala
 50 55 60

25

Glu Arg Glu Asn Pro Pro Asp Ala Tyr Val Ile Ala Cys Phe Leu Asp
 65 70 75 80

Pro Gly Leu Asp Ala Val Lys Glu Leu Thr Asp Arg Pro Val Val Gly
 85 90 95

30

Val Ala Glu Ala Ala Ile His Met Ser Ser Phe Val Ala Ala Thr Phe
 100 105 110

Ser Ile Val Ser Ile Leu Pro Arg Val Arg Lys His Leu His Glu Leu
 115 120 125

35

Val Arg Gln Ala Gly Ala Thr Asn Arg Leu Ala Ser Ile Lys Leu Pro
 130 135 140

Asn Leu Gly Val Met Ala Phe His Glu Asp Glu His Ala Ala Leu Glu
 145 150 155 160

Thr Leu Lys Gln Ala Ala Lys Glu Ala Val Gln Glu Asp Gly Ala Glu
 165 170 175

45

Ser Ile Val Leu Gly Cys Ala Gly Met Val Gly Phe Ala Arg Gln Leu
 180 185 190

Ser Asp Glu Leu Gly Val Pro Val Ile Asp Pro Val Glu Ala Ala Cys
 195 200 205

50

Arg Val Ala Glu Ser Leu Val Ala Leu Gly Tyr Gln Thr Ser Lys Ala
 210 215 220

55

Asn Ser Tyr Gln Lys Pro Thr Glu Lys Gln Tyr Leu
 225 230 235

<210> 11

<211> 25
 <212> DNA
 <213> Künstliche Sequenz

5 <220>
 <223> Beschreibung der künstlichen Sequenz: Primer5

10 <400> 11
 gccgcaagga atggtgcatg catcg 25

15 <210> 12
 <211> 30
 <212> DNA
 <213> Künstliche Sequenz

0 <220>
 <223> Beschreibung der künstlichen Sequenz: Primer6
 <400> 12
 ggtaggtgg gtccaccgag ctactgccgc 30

25 <210> 13
 <211> 5777
 <212> DNA
 <213> Künstliche Sequenz

30 <220>
 <223> Beschreibung der künstlichen Sequenz: Plasmid pOM21

<400> 13
 aattcttaag aaggagatat acatatgaga atcctcgtga tcaaccccaa cagttccagc 60
 35 gcccttactg aatcggttgc ggacgcagca caacaagttg tcgcgaccgg caccataatt 120
 tctgccatca acccctccag aggaccggcc gtcattgaag gcagctttga cgaagcactg 180
 gccacgttcc atctcattga agaggtggag cgcgctgagc gggaaaaccc gcccgacgcc 240
 tacgtcatcg catgtttcgg gcatccggga cttgacgcgg tcaaggagct gactgacagg 300
 ccagtggtag gagttgccga agctgcaatc cacatgtctt cattcgtcgc ggccaccttc 360
 45 tccattgtca gcacccctcc gagggtcagg aaacatctgc acgaactggc acggcaagcg 420
 ggggagcaga atgcctcgc ctccatcaag ctcccaaate tgggggtgat ggccttccat 480
 50 gaggacgaac atgccgcact ggagacgctc aaacaagccg ccaaggaggc ggtccaggag 540
 gacggcgccg agtcgatagt gctcggatgc gccggcatgg tgggggttgc gcgtcaactg 600
 agcgacgaac tcggcgctcc tgtcatcgac cccgtcgagg cagcttgccg cgtggccgag 660
 55 agtttggtcg ctctgggcta ccagaccagc aaagcgaact cgtatcaaaa accgacagag 720
 aagcagtacc tctagctgca gccaaagctc tggtttggcg gatgagagaa gattttcagc 780
 ctgatacaga ttaaatacaga acgcagaagc ggtctgataa aacagaattt gcctggcggc 840

5 agtagcgcg tggtcccacc tgaccccatg ccgaactcag aagtgaacg ccgtagcgcc 900
 gatggtagtg tggggtctcc ccatgcgaga gtaggggaact gccaggcatc aaataaaacg 960
 aaaggctcag tcgaaagact gggcctttcg ttttatctgt tgtttgtcgg tgaacgctct 1020
 cctgagtagg acaaatccgc cgggagcgga tttgaacgtt gcgaagcaac ggcccggagg 1080
 10 gtggcgggca ggacgcccgc cataaactgc caggcatcaa attaagcaga aggccatcct 1140
 gacggatggc ctttttgctt ttctacaaac tcttttgttt atttttctaa atacattcaa 1200
 15 atatgtatcc gctcatgaga caataaccct gataaatgct tcaataatat cgtccattcc 1260
 gacagcatcg ccagtcacta tggcgtgctg ctagcgctat atgcgttgat gcaatttcta 1320
 tgcgaccccg ttctcggagc actgtccgac cgctttggcc gccgcccagt cctgctcgt 1380
 20 tcgctacttg gagccactat cgactacgcg atcatggcga ccacaccgct cctgtggatc 1440
 ctctacgceg gacgcatcgt ggccggcatc accggcgcca cagggtgcggg tgctggcgcc 1500
 25 tatatcgccg acatcaccga tggggaagat cgggctcgcc acttcgggct catgagcgct 1560
 tgtttcggcg tgggtatggg ggcaggcccc gtggccgggg gactgttggg cgccatctcc 1620
 ttgcatgcac cattccttgc ggccggcggtg ctcaacggcc tcaacctact actgggctgc 1680
 30 ttccctaatgc aggagtcgca taaggagag cgtcgaccga tgccttgag agccttcaac 1740
 ccagtcagct ccttcgggtg ggcggggggc atgactatcg tcgcccact tatgactgtc 1800
 35 ttctttatca tgcaactcgt aggacagggt cgggcagcgc tctgggtcat ttccggcgag 1860
 gaccgcttcc gctggagcgc gacgatgatc ggctgtcgc ttgcggtatt cggaatcttg 1920
 cagccctcg ctcaagcctt cgtcactggg cccgccacca aacgtttcgg cgagaagcag 1980
 gccattatcg ccggcatggc ggccgacgcg ctgggctacg tcttgctggc gttcgcgacg 2040
 cgaggctgga tggccttccc cattatgatt cttctcgctt ccggcgccat cgggatgcc 2100
 45 gcgttgacgg ccatgctgtc caggcaggta gatgacgacc atcagggaca gcttcaagga 2160
 tcgctcgcg ctcttaccag cctaacttcg atcactggac cgctgatcgt cacggcgatt 2220
 tatgccgct cggcgagcac atggaacggg ttggcatgga ttgtaggcgc cgccctatac 2280
 50 cttgtctgcc tcccccggtt gcgtcgcggt gcatggagcc gggccacctc gacctgaatg 2340
 gaagccggcg gcacctcgct aacggattca ccactccaag aattggagcc aatcaattct 2400
 55 tgccggagaac tgtgaatgcg caaaccaacc cttggcagaa catatccatc gcgtccgcca 2460
 tctccagcag ccgcacggcg cgcattctcg gcagcggttg gtctggcca cgggtgcgca 2520
 tgatcggtg cctgtcggtt aggaccggc taggctggcg gggttgcctt actgggttagc 2580

agaatgaatc accgatacgc gagcgaacgt gaagcgactg ctgctgcaaa acgtctgcga 2640
 cctgagcaac aacatgaatg gtcttcgggt tccgtgtttc gtaaagtctg gaaacgcgga 2700
 5 agtccccctac gtgctgctga agttgcccgc aacagagagt ggaaccaacc ggtgatacca 2760
 cgatactatg actgagagtc aacgccatga gcggcctcat ttcttattct gagttacaac 2820
 10 agtccgcacc gctgtccggt agtccttcc ggtgggcgcg gggcatgact atcgtcgccg 2880
 cacttatgac tgtcttcttt atcatgcaac tcgtaggaca ggtgccggca gcgccaaca 2940
 gtcccccggc cagggggcct gccaccatac ccacgccga acaagcgccc tgcaccatta 3000
 15 tgttccggat ctgcatcgca ggatgctgct ggctaccctg tggaacacct acatctgtat 3060
 taacgaagcg ctaaccgttt ttatcaggct ctgggaggca gaataaatga tcatatcgtc 3120
 aattattacc tccacgggga gagcctgagc aaactggcct caggcatttg agaagcacac 3180
 ggtcacactg cttccggtag tcaataaacc ggtaaaccag caatagacat aagcggctat 3240
 ttaacgaccc tgccctgaac cgacgaccgg gtogaatttg ctttcgaatt tctgccattc 3300
 25 atccgcttat tatcacttat tcaggcgtag caccaggcgt ttaagggcac caataactgc 3360
 cttaaaaaaa ttacgccccg ccctgccact catcgagta ctggtgtaat tcattaagca 3420
 ttctgccgac atggaagcca tcacagacgg catgatgaac ctgaatcgcc agcggcatca 3480
 30 gcaccttgte gccttgcgta taatatttgc ccatggtgaa aacgggggcg aagaagttgt 3540
 ccatattggc cacgtttaa tcaaaactgg tgaaactcac ccagggattg gctgagacga 3600
 35 aaaacatatt ctcaataaac ctttaggga aataggccag gttttcaccg taacacgcca 3660
 catcttgcca atatatgtgt agaaactgcc ggaaatcgte gtggtattca ctccagagcg 3720
 atgaaaacgt ttcagtttgc tcatggaaaa cgggtgaaca aggggtgaaca ctatcccata 3780
 tcaccagctc accgtcttcc attgccatac gaattccgga tgagcattca tcaggcgggc 3840
 aagaatgtga ataaaggccg gataaaactt gtgcttattt ttctttacgg tctttaaaaa 3900
 45 ggccgtaata tccagctgaa cggctcggtt ataggtagat tgagcaactg actgaaatgc 3960
 ctcaaaatgt tctttacgat gccattggga tatatcaacg gtggtatatc cagtgtttt 4020
 tttctccatt ttagcttcc tagctcctga aaatctcgat aactcaaaaa atacgcccgg 4080
 50 tagtgatctt atttcattat ggtgaaagt ggaacctctt acgtgccgat caacgtctca 4140
 ttttcgcaa aagttggccc agggcttccc ggtatcaaca gggacaccag gatttattta 4200
 55 ttctgcgaag tgatcttccg tcacaggtat ttattcggcg caaagtgcgt cgggtgatgc 4260
 tgccaaacta ctgatttagt gtatgatggt gtttttgagg tgctccagt gcttctgttt 4320
 ctatcagctg tccctcctgt tcagctactg acgggggtgt gcgtaacggc aaaagcaccg 4380

5 cccgacatca gcgctagcgg agtgtatact ggcttactat gttggcactg atgaggggtgt 4440
 cagtgaagtg cttcatgtgg caggagaaaa aaggctgcac cgggtgcgtca gcagaatatg 4500
 tgatacagga tatattccgc ttcctcgtc actgactcgc tacgctcggg cgttcgactg 4560
 cggcgagcgg aaatggctta cgaacggggc ggagatttcc tgggaagatgc caggaagata 4620
 10 cttaacaggg aagtgagagg gccgcggcaa agccgttttt ccataggctc cgtccccctg 4680
 acaagcatca cgaaatctga cgctcaaadc agtggtggcg aaacccgaca ggactataaa 4740
 15 gataccaggc gtttccccctg gcggctccct cgtgcgctct cctgttccctg cttttcgggt 4800
 taccgggtgc attccgctgt tatggccgcg tttgtctcat tccacgcctg acactcagtt 4860
 cccggtaggc agttcgtcc aagctggact gtatgcacga accccccgtt cagtccgacc 4920
 20 gctgcgcctt atccggtaac tatcgtcttg agtccaaccc ggaaagacat gcaaaagcac 4980
 cactggcagc agccactggg aattgattta gaggagttag tcttgaagtc atgcgccggg 5040
 taaggctaaa ctgaaaggac aagttttggg gactgcgtc ctccaagcca gttacctcgg 5100
 25 ttcaaagagt tggtagctca gagaaccttc gaaaaaccgc cctgcaaggc gggttttttcg 5160
 ttttcagagc aagagattac gcgcagacca aaacgatctc aagaagatca tcttattaat 5220
 30 cagataaaat atttcaagat ttcagtgcga tttatctctt caaatgtagc acctgaagtc 5280
 agccccatac gatataagtt gtaattctca tggttgacag cttatcatcg ataagcttta 5340
 atgcggtagt ttatcacagt taaattgcta acgcagtcag gcaccgtgta tgaaatctaa 5400
 35 caatgcgtc atcgtcatcc tcggcacctg caccctggat gctgtaggca taggcttggg 5460
 tatgccggta ctgccgggccc tcttgccgga ttagtcatgc cccgcgccc cccggaaggag 5520
 ctgactgggt tgaaggctct caagggcatc ggtcgacgct ctcccttatg cgactcctgc 5580
 attaggaagc agcccagtag taggttgagg ccgttgagca ccgcccgcgc aaggaatggg 5640
 45 gcatgcatcg atcaccacaa ttcagcaaat tgtgaacatc atcacgttca tctttccctg 5700
 gttgccaatg gccatttttc ctgtcagtaa cgagaaggtc gcgaattcag gcgcttttta 5760
 gactggtcgt aatgaac 5777

50

<210> 14

<211> 7175

<212> DNA

<213> Künstliche Sequenz

55

<220>

<223> Beschreibung der künstlichen Sequenz: Plasmid pOM22

<400> 14

aattcttaag aaggagatat acatatgacc ctgcagaaag cgcaagcgna gcgcattgag 60
aaagagatct gggagctctc ccggttctcg gcggaaggcc ccggtgttac ccggctgacc 120
5 tacactccag agcatgccgc cgcgcgggaa acgctcattg cggctatgga agcggccgct 180
ttgagcgttc gtgaagacgc tctcgggaac atcatcggcc gacgtgaagg cactgatccg 240
cagctccctg cgatcgcggt cggttcacac ttcgattctg tccgaaacgg cgggatgttc 300
10 gatggcactg caggcgtggt gtgcgccctt gaggctgccc gggatgatgt ggagagcggc 360
tacgtgaatc ggcattccatt tgagttcatc gcgatcgtgg aggaggaagg ggcccgttc 420
15 agcagtggca tgttgggcgg ccgggccatt gcaggtttgg tcgccgacag ggaactggac 480
tctttggttg atgaggatgg agtgctcggt aggcaggcgg ctactgcctt cggcttgaag 540
cggggcgaac tgcaggctgc agcccgtcc gcggcggacc tgcgtgcttt tatcgaacta 600
cacattgaac aaggaccgat cctcgagcag gagcaaatag agatcggagt tgtgacctcc 660
atcgttggcg ttcgcgcatt gcgggttgc gtcaaaggca gaagcgcaca cgccggcaca 720
25 acccccatgc acctgcgcca ggatgcgctg gtaccgcgcg ctctcatggt gcgggaggtc 780
aaccggttcg tcaacgagat cgccgatggc acagtggcta ccgttggcca cctcacagtg 840
gccccgggtg gcggcaacca ggtcccgggg gaggtggagt tcacactgga cctgcgttct 900
30 ccgcatgagg agtcgctccg ggtgttgatc aaccgcatct cggatcatggt cggcgaggtc 960
gcctcgcagg ccggtgtggc tgccgatgtg gatgaatttt tcaatctcag cccggtgcag 1020
35 ctggctccta ccatggtgga cgccgttcgc gaagcggcct cggccctgca gttcacgcac 1080
cgggatatca gcagtggggc gggccacgac tcgatgttca tcgccaggt cacggacgtc 1140
ggaatggttt tcgttccaag ccgtgctggc cggagccacg tcccgaaga atggaccgat 1200
ttcgatgacc ttcgcaaggg aactgaggtt gtccctccgg taatgaaggc acttgaccgg 1260
ggatcccatc atcatcatca tcattgactg .cagccaagct tctgttttgg cggatgagag 1320
45 aagattttca gcctgataca gattaaatca gaacgcagaa gcggtctgat aaaacagaat 1380
ttgcctggcg gcagtagcgc ggtggtccca cctgacccca tgccgaactc agaagtgaaa 1440
cgccgtagcg ccgatggtag tgtggggtct ccccatgcga gagtagggaa ctgccaggca 1500
50 tcaaataaaa cgaaaggctc agtcgaaaga ctgggccttt cgttttatct gttgtttgtc 1560
ggtgaacgt ctctgagta ggacaaatcc gccgggagcg gatttgaacg ttgcgaagca 1620
55 acggccccga ggggtggcggg caggacgcc gccataaact gccaggcatc aaattaagca 1680
gaaggccatc ctgacggatg gcctttttgc gtttctacaa actcttttgt ttatttttct 1740
aaatacatc aaatatgtat ccgctcatga gacaataacc ctgataaatg cttcaataat 1800

attgaaaaag gaagagtatg agtattcaac atttccgtgt cgcccttatt cccttttttg 1860
 5 cggcattttg ccttcctgtt ttgctcacc cagaaacgct ggtgaaagta aaagatgctg 1920
 aagatcagtt ggggtgcacga gtgggttaca tcgaactgga tctcaacagc ggtaagatcc 1980
 ttgagagttt tcgccccgaa gaacgttttc caatgatgag cactttttaa gttctgctat 2040
 10 gtggcgcggt attatcccgt gttgacgccg ggcaagagca actcggtcgc cgcatacact 2100
 attctcagaa tgacttgggt gagtactcac cagtcacaga aaagcatctt acggatggca 2160
 15 tgacagtaag agaattatgc agtgctgcca taaccatgag tgataacact ggggccaact 2220
 tactttctgac aacgatcggg ggaccgaagg agctaaccgc ttttttgac aacatggggg 2280
 atcatgtaac tcgccttgat cgttgggaac cggagctgaa tgaagccata ccaaacgacg 2340
 20 agcgtgacac cacgatgcct gtagcaatgg caacaacggt gcgcaaacta ttaactggcg 2400
 aactacttac tctagcttcc cggcaacaat taatagactg gatggaggcg gataaagttg 2460
 25 caggaccact tctgcgctcg gcccttccgg ctggctggtt tattgctgat aaatctggag 2520
 ccggtgagcg tgggtctcgc ggtatcattg cagcactggg gccagatggt aagccctccc 2580
 gtatcgtagt tatctacacg acggggagtc aggcaactat ggatgaacga aatagacaga 2640
 30 tcgctgagat aggtgcctca ctgattaagc attggttaact gtcagaccaa gtttactcat 2700
 atatacttta gattgattta aaacttcatt ttaatttaa aaggatctag gtgaagatcc 2760
 35 tttttgataa tctcatgacc aaaatccctt aacgtgagtt ttcgttccac tgagcgtcag 2820
 accccgtaga aaagatcaaa ggatcttctt gagatccttt ttttctgcgc gtaatctgct 2880
 gcttgcaaac aaaaaaacca ccgctaccag cgggtggttg tttgccggat caagagctac 2940
 caactctttt tccgaaggta actggcttca gcagagcgca gataccaaat actgtccttc 3000
 tagtgtagcc gtagttaggc caccacttca agaactctgt agcaccgcct acatacctcg 3060
 45 ctctgctaata cctgttacca gtggctgctg ccagtggcga taagtcgtgt cttaccgggt 3120
 tggactcaag acgatagtta ccggataagg cgcagcggtc gggctgaacg gggggttcgt 3180
 gcacacagcc cagcttggag cgaacgacct acaccgaact gagataccta cagcgtgagc 3240
 50 tatgagaaag cgccacgctt cccgaaggga gaaaggcgga caggtatccg gtaagcggca 3300
 gggtcggaac aggagagcgc acgagggagc ttccaggggg aaacgcctgg tatctttata 3360
 55 gtctgtcgg gtttcgccac ctctgacttg agcgtcgatt tttgtgatgc tcgtcagggg 3420
 ggcggagcct atggaaaaac gccagcaacg cggccttttt acggttctcg gccttttgct 3480
 ggccttttgc tcacatgttc tttctgctg tatccctga ttctgtggat aaccgtatta 3540

ccgcctttga gtgagctgat accgctcgcc gcagccgaac gaccgagcgc agcgagtcag 3600
tgagcgagga agcggaagag cgctgatgc ggtattttct ccttacgcat ctgtgcggta 3660
5 tttcacaccg catatatggt gactctcag tacaatctgc tctgatgccg catagttaag 3720
ccagtataca ctccgctatc gctacgtgac tgggtcatgg ctgcgccccg acacccgccca 3780
acacccgctg acgcgccttg acgggcttgt ctgctcccgg catccgctta cagacaagct 3840
10 gtgaccgtct ccgggagctg catgtgtcag aggttttcac cgtcatcacc gaaacgcgcg 3900
aggcagctgc ggtaaagctc atcagcgtgg tcgtgaagcg attcacagat gtctgcctgt 3960
15 tcatccgctg ccagctcggt gagtttctcc agaagcgcta atgtctggct tctgataaag 4020
cgggccatgt taagggcggt tttttcctgt ttggtcactt gatgcctccg tgtaaggggg 4080
aattttctgtt catgggggta atgataccga tgaacgaga gaggatgctc acgatacggg 4140
ttactgatga tgaacatgcc cggttactgg aacgttgtga gggtaaaca ctggcggtat 4200
ggatgcggcg ggaccagaga aaaatcactc aggggtcaatg ccagcgcttc gttatacag 4260
25 atgtaggtgt tccacagggt agccagcagc atcctgcgat gcagatccgg aacataatgg 4320
tgcagggcgc tgacttccgc gtttcagac ttacgaaac acggaaaccg aagaccattc 4380
atgttggtgc tcaggtcgca gacgttttgc agcagcagtc gcttcacgtt cgctcgcgta 4440
30 tcggtgattc attctgctaa ccagtaaggc aaccccgcca gcctagccgg gtctcaacg 4500
acaggagcac gatcatgcgc acccgtggcc aggacccaac gctgcccag atgcgccgcg 4560
35 tgcggctgct ggagatggcg gacgcgatgg atatgttctg ccaaggggtg gtttgcgcat 4620
tcacagttct ccgcaagaat tgattggctc caattcttgg agtggatgaat ccgttagcga 4680
ggtgccgccg gcttccattc aggtcgaggt ggcccggctc catgcaccgc gacgcaaccg 4740
ggggaggcag acaaggata gggcgccgc tacaatccat gccaacccgt tccatgtgct 4800
cgccgaggcg gcataaatcg ccgtgacgat cagcgggtcca gtgatcgaag ttaggctggt 4860
45 aagagccgcg agcgatcctt gaagctgtcc ctgatggtcg tcatctacct gcctggacag 4920
catggcctgc aacgcgggca tcccgatgcc gccggaagcg agaagaatca taatggggaa 4980
ggccatccag cctcgcgctg cgaacgccag caagacgtag ccagcgcggt cggccgccat 5040
50 gccggcgata atggcctgct tctcgccgaa acgtttggtg gcgggaccag tgacgaaggc 5100
ttgagcgagg gcgtgcaaga ttccgaatac cgcaagcgac aggccgatca tcgtcgcgct 5160
55 ccagcgaaag cggtcctcgc cgaaaatgac ccagagcgct gccggcacct gtcctacgag 5220
ttgcatgata aagaagacag tcataagtgc ggcgacgata gtcatgcccc gcgcccaccg 5280
gaaggagctg actgggttga aggtcttcaa gggcatcggt cgacgctctc ccttatgcga 5340

5 ctctgcatt aggaagcagc ccagtagtag gttgaggccg ttgagcaccg ccgccgcaag 5400
 gaatggtgca tgcacgcatc accacaattc agcaaattgt gaacatcatc acgttcatct 5460
 10 ttccctgggt gccaatggcc ctttttctg tcagtaacga gaaggtcgcg aattcaggcg 5520
 ctttttagac tggtcgtaat gaacaattct taagaaggag atatacatat gtttgacgta 5580
 atagttaaga actgccgtat ggtgtccagc gacggaatca ccgaggcaga cattctggtg 5640
 aaagacggca aagtcgccgc aatcagctcg gacacaagt atgttgaggc gagccgaacc 5700
 15 attgacgcgg gtggcaagtt cgtgatgccg ggcgtggctg atgaacatgt gcatacatc 5760
 gacatggatc tgaagaaccg gtatggccgc ttcgaactcg attccgagtc tgcggccgtg 5820
 ggaggcatca ccaccatctt tgagatgccg tttaccttc cgcccaccac cactttggac 5880
 20 gccttctctg aaaagaagaa gcaggcgggg cagcggttga aagttgactt cgcgctctat 5940
 ggcggtggag tgccgggaaa cctgcccgag atccgcaaaa tgcacgacgc cggcgagtg 6000
 25 ggcttcaagt caatgatggc agcctcagtt ccgggcatgt tcgacgccgt cagcgacggc 6060
 gaactgttcg aaatcttcca ggagatcgca gcctgtggtt cagtcgccgt ggtccatgcc 6120
 gagaatgaaa cgatcattca agcgtccag aagcagatca aagccgctgg tcgcaaggac 6180
 30 atggccgcct acgaggcatc ccaaccagtt ttccaggaga acgaggccat tcagcgtgcg 6240
 ttactactgc agaaagaagc cggctgtcga ctgattgtgc ttcacgtgag caaccctgac 6300
 35 ggggtcgagc tgatacatc ggcgcaatcc gagggccagg acgtccactg cgagtcgggt 6360
 ccgcagtatc tgaatatcac cacggacgac gccgaacgaa tcggaccgta tatgaaggtc 6420
 gcgccgccg tccgctcagc cgagatgaac gtcagattat gggaacaact tgagaacggg 6480
 40 ctcatcgaca cccttgggtc agaccacggc ggacatcctg tcgaggacaa agaaccggc 6540
 tggaaggacg tgtggaaagc cggcaacggt gcgctgggcc ttgagacatc cctgcctatg 6600
 45 atgctgacca acggagtga taaaggcagg ctatccttgg aacgcctcgt cgaggtgatg 6660
 tgcgagaaac ctgcgaagct ctttggcatc tatccgcaga agggcacgct acaggttggg 6720
 tccgacgccg atctgctcat cctcgatctg gatattgaca ccaaagtgga tgcctcgag 6780
 50 ttccgatccc tgcataagta cagcccgttc gacgggatgc ccgtcacggg tgcaccggtt 6840
 ctgacgatgg tgcgcggaac ggtggtggca gagaaggag aagttctggt cgagcaggga 6900
 55 ttcgccagct tcgtcaccgc tcacgactac gaggcgtcga agtgaggatc tcgacgctct 6960
 cccttatgcg actcctgcat taggaagcag ccagtagta ggttgaggcc gttgagcacc 7020
 gccgccgcaa ggaatggtgc atgcatcgat caccacaatt cagcaaattg tgaacatcat 7080

cacgttcate tttccctggg tgccaatggc ccattttcct gtcagtaacg agaagggtcgc 7140
 gaattcaggc gctttttaga ctgggtcgtaa tgaac 7175

5

<210> 15
 <211> 5989
 <212> DNA
 <213> Künstliche Sequenz

10

<220>
 <223> Beschreibung der künstlichen Sequenz: Plasmid pDHYH

<400> 15
 15 aattcttaag aaggagatat acatatggat gcaaagctac tggttggcgg cactattggt 60
 tctctgaccg gcaaaatccg agccgacgtg ctgattgaaa acggcaaagt cgccgctgtc 120
 20 ggcattgctgg acgcccgcgac gccggacaca gttgagcggg ttgactgcga cggcaaatac 180
 gtcattgccc gcggtatcga cgttcacacc cacatcgact cccccctcat ggggaccacc 240
 accgccgatg attttgtcag cggaacgatt gcagccgcta ccggcgggaac aacgaccatc 300
 25 gtgatattcg gacagcagct cgccggcaag aacctgctgg aatccgcaga cgcgaccac 360
 aaaaaggcgc aggggaaatc cgtcattgat tacggcttcc atatgtgcgt gacgaacctc 420
 30 tatgacaatt tcgattccca tatggcagaa ctgacacagg acggaatctc cagtttcaag 480
 gtcttcatgg cctaccgagg aagcctgatg atcaacgacg gcgaactggt cgacatcctc 540
 aaggggagtgc gctccagcgg tgccaaaacta tgcgtccacg cagagaacgg cgacgtcatc 600
 35 gacaggatcg ccgccgacct ctacgcccac ggaaaaaccg ggcccgggac ccacgagatc 660
 gcacgcccgc cggaatcgga agtcgaagca gtcagccggg ccatcaagat ctcccggatg 720
 gccgaggtgc cgctgtatct cgtgcatctt tccaccagg gggccgtcga ggaagtagct 780
 gccgcgcaga tgacaggatg gccaatcagc gccgaaacgt gcacccacta cctgtcgtg 840
 agccggggaca tctacgacca gccgggattc gagccggcca aagctgtcct cacaccaccg 900
 45 ctgcgcacac aggaacacca ggacgcgttg tggagaggca ttaacaccgg tgcgtcagc 960
 gtcgtcagtt ccgaccactg ccccttctgc tttgaggaaa agcagcggat gggggcagat 1020
 gacttccggc agatcccca cggcggggccc ggcggtggagc accgaatgct cgtgatgtat 1080
 50 gagaccggtg tcgcggaagg aaaaatgacg atcgagaaat tcgtcgaggt gactgccgag 1140
 aaccgggcca agcaattcga tatgtacctg aaaaagggaa caattgcacc gggctccgat 1200
 55 gcagacatca tcgtggtcga ccccaacgga acaaccctca tcagtgccga caccacaaaa 1260
 caaaacatgg actacacgct gttcgaaggc ttcaaaatcc gttgctccat cgaccagggtg 1320
 ttctcgcgtg gcgacctgat cagcgtcaaa ggccgaatat tcggcaccgc cgcccgcggc 1380

gaattcatca agcggagcgc ttggagccac ccgcagttcg aaaaataaaa gcttggctgt 1440
5 tttggcggat gagagaagat tttcagcctg atacagatta aatcagaacg cagaagcggg 1500
ctgataaaac agaatttgcc tggcggcagt agcgcgggtg tcccacctga ccccatgccg 1560
aactcagaag tgaaacgccg tagcgccgat ggtagtgtgg ggtctcccca tgcgagagta 1620
10 gggaactgcc aggcatacaa taaaacgaaa ggctcagtcg aaagactggg cctttcgttt 1680
tatctgttgt ttgtcggatga acgctctcct gagtaggaca aatccgccgg gagcggattt 1740
15 gaacgttgcg aagcaacggc ccggaggggtg gcgggcagga cgcccgccat aaactgccag 1800
gcatcaaatt aagcagaagg ccactctgac ggatggcctt tttgcgtttc tacaaactct 1860
tttgtttatt tttctaaata cattcaaata tgtatccgct catgagacaa taaccctgat 1920
20 aaatgcttca ataatttga aaaaggaaga gtatgagtat tcaacatttc cgtgtcgccc 1980
ttattccctt ttttgccgca ttttgccttc ctgtttttgc tcacccagaa acgctgggtga 2040
aagtaaaaaga tgctgaagat cagttgggtg cagcagtggtg ttacatcgaa ctggatctca 2100
25 acagcggtaa gatccttgag agttttcgcc ccgaagaacg ttttccaatg atgagcactt 2160
ttaaagttct gctatgtggc gcggtattat cccgtgttga cgccgggcaa gagcaactcg 2220
30 gtcgccgcat acactattct cagaatgact tgggtgagta ctcaccagtc acagaaaagc 2280
atcttacgga tggcatgaca gtaagagaat tatgcagtgc tgccataacc atgagtata 2340
35 acactgcggc caacttactt ctgacaacga tcggaggacc gaaggagcta accgcttttt 2400
tgcacaacat gggggatcat gtaactcgcc ttgatcggtg ggaaccggag ctgaatgaag 2460
ccataccaaa cgacgagcgt gacaccacga tgctgtagc aatggcaaca acgttgcgca 2520
aactattaac tggcgaacta cttactctag cttcccggca acaattaata gactggatgg 2580
aggcggataa agttgcagga ccacttctgc gctcggccct tccggctggc tggtttattg 2640
45 ctgataaatc tggagccggg gagcgtgggt ctcgcggtat cattgcagca ctggggccag 2700
atggtaagcc ctcccgatc gtagttatct acacgacggg gagtcaggca actatggatg 2760
aacgaaatag acagatcgct gagatagggt cctcactgat taagcattgg taactgtcag 2820
50 accaagttta ctcatatata ctttagattg atttaaaact tcatttttaa tttaaaagga 2880
tctaggtgaa gatccttttt gataatctca tgacaaaat cccttaacgt gagttttcgt 2940
tccactgagc gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat cttttttttc 3000
55 tgcgcgtaat ctgctgctg caaacaaaaa aaccaccgct accagcgggt gtttgtttgc 3060
cggatcaaga gctaccaact ctttttccga aggttaactgg cttcagcaga gcgcagatac 3120

caaatactgt ccttctagt tagccgtagt taggccacca cttcaagaac tctgtagcac 3180
 cgctacata cctcgctctg ctaatcctgt taccagtggc tgctgccagt ggcgataagt 3240
 5 cgtgtcttac cgggttgac tcaagacgat agttaccgga taaggcgag cggtcgggct 3300
 gaacgggggg ttcgtgcaca cagcccagct tggagcgaac gacctacacc gaactgagat 3360
 10 acctacagcg tgagctatga gaaagcgcca cgcttcccga agggagaaag gcggacaggt 3420
 atccggttag cggcagggtc ggaacaggag agcgcacgag ggagcttcca gggggaaacg 3480
 cctggtatct ttatagtct gtcgggtttc gccacctctg acttgagcgt cgatttttgt 3540
 15 gatgctcgtc aggggggagg agcctatgga aaaacgccag caacgcggcc tttttacggt 3600
 tcttggcctt ttgctggcct tttgctcaca tggtctttcc tgcgttatcc cctgattctg 3660
 20 tggataaccg tattaccgcc tttgagttag ctgataccgc tcgccgcagc cgaacgaccg 3720
 agcgcagcga gtcagttagc gaggaagcgg aagagcgct gatgcggtat tttctcctta 3780
 cgcatctgtg cggattttca caccgcatat atggtgcact ctgagtaca tctgctctga 3840
 25 tgccgcatag ttaagccagt atacactccg ctatcgctac gtgactgggt catggctgcg 3900
 ccccgacacc cgccaacacc cgctgacgag ccctgacggg cttgtctgct cccggcatcc 3960
 30 gcttacagac aagctgtgac cgtctccggg agctgcatgt gtcagagggt ttcaccgtca 4020
 tcaccgaaac gcgcgaggca gctgcggtaa agctcatcag cgtggctcgtg aagcgattca 4080
 cagatgtctg cctgttcac cgcgtccagc tcgttgagtt tctccagaag cgttaatgtc 4140
 35 tggcttctga taaagcgggc catgttaagg gcggtttttt cctgtttggt cacttgatgc 4200
 ctccgtgtaa gggggaattt ctgttcattg gggtaatgat accgatgaaa cgagagagga 4260
 40 tgctcacgat acgggttact gatgatgaac atgcccgggt actggaacgt tgtgagggta 4320
 aacaactggc ggtatggatg cggcgggacc agagaaaaat cactcagggt caatgccagc 4380
 gcttcgttaa tacagatgta ggtgttcac agggtagcca gcagcatcct gcgatgcaga 4440
 45 tccggaacat aatggtgcag ggcgctgact tccgcgtttc cagactttac gaaacacgga 4500
 aaccgaagac cattcatggt gttgctcagg tcgcagacgt tttgcagcag cagtcgcttc 4560
 50 acgttcgctc gcgtatcggg gattcattct gctaaccagt aaggcaaccc cgccagccta 4620
 gccgggtcct caacgacagg agcacgatca tgcgcacccg tggccaggac ccaacgctgc 4680
 ccgagatgcg ccgcgtgcgg ctgctggaga tggcggacgc gatggatatg ttctgccaa 4740
 55 ggttggtttg cgcattcaca gttctccgca agaattgatt ggctccaatt cttggagtgg 4800
 tgaatccgtt agcgagggtc cgccggcttc cattcagggt gaggtggccc ggctccatgc 4860
 accgcgacgc aacgcgggga ggcagacaag gtatagggcg gcgcctacaa tccatgccaa 4920

5 cccgttccat gtgctcgccg aggcggcata aatcgccgtg acgatcagcg gtccagtgat 4980
 cgaagttagg ctggtaagag ccgcgagcga tccttgaagc tgtccctgat ggctgctcgc 5040
 tacctgcctg gacagcatgg cctgcaacgc gggcatcccc atgccgcggg aagcgagaag 5100
 aatcataatg gggaaggcca tccagcctcg cgctcgcaac gccagcaaga cgtagcccag 5160
 10 cgcgtcgccc gccatgccgg cgataatggc ctgcttctcg ccgaaacggt tgggtggcggg 5220
 accagtgcg aaggcttgag cgagggcggt caagattccg aataccgcaa gcgacaggcc 5280
 15 gatcatcgct gcgctccagc gaaagcgggt ctcgccgaaa atgaccaga gcgctgccgg 5340
 cacctgtcct acgagttgca tgataaagaa gacagtcata agtgccggcg cgatagtcac 5400
 gccccgcgcc caccggaagg agctgactgg gttgaaggct ctcaagggca tcggtcgacg 5460
 20 ctctccctta tgcgactcct gcattaggaa gcagcccagt agtaggttga ggccgttgag 5520
 caccgcgcgc gcaaggaatg gtgcatgctc gatggctacg agggcagaca gtaagtggat 5580
 ttaccataat cccttaattg tacgcaccgc taaaacgcgt tcagcgcgat cacggcagca 5640
 25 gacaggtaaa aatggcaaca aaccacccta aaaactgcgc gatcgcgctt gataaatttt 5700
 aaccgtatga atacctatgc aaccagaggg tacaggccac attaccccca cttaatccac 5760
 30 tgaagctgcc atttttcatg gtttcacat cccagcgaag ggccatgcat gcacgaaaat 5820
 taatacgacg aaattaatac gactcactat agggcaattg cgatcaccac aattcagcaa 5880
 attgtgaaca tcatcacggt catctttccc tgggtgcaa tggcccatth tcctgtcagt 5940
 35 aacgagaagg tcgcgaattc aggcgctttt tagactgggc gtaatgaac 5989

<210> 16
 <211> 6958
 <212> DNA
 <213> Künstliche Sequenz

45 <220>
 <223> Beschreibung der künstlichen Sequenz: Plasmid
 pJAVIER16

<400> 16
 50 aattcttaag aaggagatat acatatggcg aaaaacttga tgctcgcggt cgctcaagtc 60
 ggccggtatcg atagttcgga atcaagaccg gaagtcgctg cccgcttgat tgccctgctg 120
 gaagaagcag ctccccaggg ccgcggaactg gtggtctttc ccgaactcac gctgaccacg 180
 55 ttcttcccgcc gtacctgggt cgaagaaggc gacttcgagg aatacttcga taaatccatg 240
 cccaatgacg acgtcgcgcc ccttttcgaa ccgcgcaaag accttggcgt gggcttctac 300
 ctcggatacg ccgaactgac cagtgatgag aagcgggtaca acacatcaat tctggtgaac 360

aagcacggcg acatcgtcgg caagtaccgc aagatgcac tgccggggcca cgccgataac 420
 5 cgggaaggac tacccaacca gcaccttgaa aagaaatact tccgcgaagg agatctcgga 480
 ttcggtgtct tcgacttcca cggcgtgcag gtcggaatgt gtctctgcaa cgaccggcga 540
 tggcggagg tctaccgctc tttggccctg cagggagcag agctcgtcgt cctgggctac 600
 10 aacacccccg atttcgttcc cggctggcag gaagagcctc acgcgaagat gttcacgcac 660
 cttctttcac ttcaggcagg ggcataccag aactcggtat ttgtggcggc tgccggcaag 720
 tcgggcttcg aagacgggca ccacatgac ggccgatcag cggtcgccgc gcccagcggc 780
 15 gaaatcctgg caaaagcagc cggcgagggc gatgaagtcg tcgttgtgaa agcagacatc 840
 gacatgggca agccctataa ggaaagcgtc ttcgacttcg ccgcccacgc gcgccccgac 900
 20 gcatacggca tcatcgccga aaggaaaggg cggggcgccc cactgcccgt cccgttcaac 960
 gtgaatgact aaggatccga aggagatata catatggatg caaagctact gggtggcggc 1020
 actattgttt cctcgaccgg caaaatccga gccgacgtgc tgattgaaaa cggcaaagtc 1080
 25 gccgctgtcg gcatgctgga cgccgcgacg ccggacacag ttgagcgggt tgactgcgac 1140
 ggcaaatacg tcatgcccg cggtatcgac gttcacaccc acatcgactc cccctcatg 1200
 30 gggaccacca ccgccgatga ttttgtcagc ggaacgattg cagccgctac cggcggaaca 1260
 acgaccatcg tcgatttcgg acagcagctc gccggcaaga acctgctgga atccgcagac 1320
 gcgcaccaca aaaaggcgca ggggaaatcc gtcattgatt acggcttcca tatgtgcgtg 1380
 35 acgaacctct atgacaattt cgattcccat atggcagaac tgacacagga cggaatctcc 1440
 agtttcaagg tcttcatggc ctaccgcgga agcctgatga tcaacgacgg cgaactgttc 1500
 gacatcctca agggagtcgg ctccagcggg gccaaactat gcgtccacgc agagaacggc 1560
 gacgtcatcg acaggatcgc cgccgacctc tacgcccag gaaaaaccgg gcccgggacc 1620
 cagagatcg cagccccgc ggaatcgga gtogaagcag tcagccgggc catcaagatc 1680
 45 tcccggatgg ccgaggtgcc gctgtatttc gtgcatcttt ccaccaggg ggccgtcgag 1740
 gaagtagctg ccgcgcagat gacaggatgg ccaatcagcg ccgaaacgtg caccactac 1800
 50 ctgtcgctga gccgggacat ctacgaccag ccgggattcg agccggccaa agctgtcctc 1860
 acaccaccgc tgcgcacaca ggaacaccag gacgcgttgt ggagaggcat taacaccggt 1920
 gcgctcagcg tcgtcagttc cgaccactgc cccttctgct ttgaggaaaa gcagcggatg 1980
 55 ggggcagatg acttccggca gatccccaac ggcggggccg gcgtggagca ccgaatgctc 2040
 gtgatgtatg agaccggtgt cgcggaagga aaaatgacga tcgagaaatt cgtcgaggtg 2100

actgccgaga acccggccaa gcaattcgat atgtacccga aaaagggaaac aattgcaccg 2160
ggctccgatg cagacatcat cgtgggacgac cccaacggaa caaccctcat cagtgccgac 2220
5 acccaaaaac aaaacatgga ctacacgctg ttcgaaggct tcaaaatccg ttgctccatc 2280
gaccaggtgt tctcgcgtgg cgacctgac agcgtcaaag gcgaatatgt cggcaccgc 2340
ggccgcggcg aattcatcaa gcgagcgct tggagccacc cgcagttcga aaaataaaaag 2400
10 cttggctgtt ttggcggatg agagaagatt ttcagcctga tacagattaa atcagaacgc 2460
agaagcggtc tgataaaaca gaatttgcct ggccggcagta gcgcgggtgg cccacctgac 2520
15 cccatgccga actcagaagt gaaacgccgt agcgcgatg gtagtggtgg gtctcccat 2580
gcgagagtag ggaactgcca ggcatacaat aaaacgaaag gctcagtcga aagactgggc 2640
ctttcgtttt atctgttgtt tgctcgtgaa cgctctcctg agtaggacaa atccgccggg 2700
agcggatttg aacgttgca agcaacggc cggagggtgg cgggcaggac gcccgccata 2760
aactgccagg catcaaatta agcagaaggc catcctgacg gatggccttt ttgcgtttct 2820
25 acaaactcct ttgtttatct ttctaaatac attcaaatat gtatccgctc atgagacaat 2880
aaccctgata aatgcttcaa taatattgaa aaaggaagag tatgagtatt caacatttcc 2940
gtgtcgccct tattcccttt ttgcgccat ttgacctcc tgtttttgct caccagaaa 3000
30 cgctgggtgaa agtaaaagat gctgaagatc agttgggtgc acgagtgggt tacatcgaac 3060
tggatctcaa cagcggtaag atccttgaga gttttcgccc cgaagaacgt tttccaatga 3120
35 tgagcacttt taaagtctg ctatgtggcg cggattatc cctgtgtgac gccgggcaag 3180
agcaactcgg tcgccgata cactattctc agaatgactt ggttgagtac tcaccagtca 3240
cagaaaagca tcttacggat ggcatacag taagagaatt atgcagtgt gccataacca 3300
tgagtataa cactgcggcc aacttacttc tgacaacgat cggaggaccg aaggagctaa 3360
ccgctttttt gcacaacatg ggggatcatg taactcgctt tgatcgttgg gaaccggagc 3420
45 tgaatgaagc cataccaaac gacgagcgtg acaccacgat gcctgtagca atggcaacaa 3480
cggtgcgcaa actattaact ggcgaactac ttactctagc ttccggcaa caattaatag 3540
actggatgga ggccgataaa gttgcaggac cacttctgcg ctccggcctt cgggtggct 3600
50 ggtttattgc tgataaatct ggagccggtg agcgtgggtc tcgcggtatc attgcagcac 3660
tggggccaga tggtaagccc tcccgatatc tagttatcta cagcagggg agtcaggcaa 3720
55 ctatggatga acgaaataga cagatcgctg agataggtgc ctactgatt aagcattggg 3780
aactgtcaga ccaagtttac tcatatatac tttagattga tttaaaactt catttttaat 3840
ttaaaaggat ctaggtgaag atcctttttg ataattctcat gacaaaatc ccttaacgtg 3900

agtttttcgtt ccactgagcg tcagaccccg tagaaaagat caaaggatct tcttgagatc 3960
ctttttttct gcgcgtaatc tgctgcttgc aaacaaaaaa accaccgcta ccagcggtagg 4020
5 tttgtttgcc ggatcaagag ctaccaactc tttttccgaa ggtaactggc ttcagcagag 4080
cgcagatacc aaatactgtc cttctagtgt agccgtagtt aggccaccac ttcaagaact 4140
10 ctgtagcacc gcctacatac ctgcctctgc taatcctggt accagtggct gctgccagtg 4200
gcgataagtc gtgtcttacc gggttggact caagacgata gttaccggat aaggcgcagc 4260
ggtcgggctg aacgggggggt tcgtgcacac agcccgactt ggagcgaacg acctacaccg 4320
15 aactgagata cctacagcgt gagctatgag aaagcgccac gcttcccgaa gggagaaagg 4380
cggacaggta tccggtaagc ggcagggctc gaacaggaga gcgcacgagg gagcttccag 4440
20 ggggaaacgc ctggtatctt tatagtcctg tcgggtttcg ccacctctga cttgagcgtc 4500
gatttttgtg atgctcgtca ggggggcgga gcctatggaa aaacgccagc aacgcggcct 4560
ttttacgggt cctggccttt tgctggcctt ttgctcacat gttctttcct gcgttatccc 4620
25 ctgattctgt ggataaccgt attaccgcct ttgagtgagc tgataccgct cgccgcagcc 4680
gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga agagcgccctg atgcggtatt 4740
30 ttctccttac gcactctgtc ggtatttcac accgcatata tgggtgcactc tcagtacaat 4800
ctgctctgat gccgcatagt taagccagta tacactccgc tatcgctacg tgactgggtc 4860
atggctgcgc cccgacaccc gccaacaccc gctgacgcgc cctgacgggc ttgtctgtc 4920
35 ccggcatccg cttacagaca agctgtgacc gtctccggga gctgcatgtg tcagaggttt 4980
tcaccgtcat caccgaaacg cgcgaggcag ctgcggtaaa gctcatcagc gtggctcgtga 5040
agcgattcac agatgtctgc ctgttcatcc gcgtccagct cgttgagttt ctccagaagc 5100
gttaatgtct ggcttctgat aaagcgggccc atgttaaggg cggttttttc ctgtttggtc 5160
acttgatgcc tccgtgtaag ggggaatttc tgttcatggg ggtaatgata ccgatgaaac 5220
45 gagagaggat gctcacgata cgggttactg atgatgaaca tgcccggta ctggaacggt 5280
gtgagggtaa acaactggcg gtatggatgc ggcgggacca gagaaaaatc actcagggtc 5340
50 aatgccagcg cttegttaat acagatgtag gtgttccaca gggtagccag cagcatcctg 5400
cgatgcagat ccggaacata atgggtgcagg gcgctgactt ccgcgtttcc agactttacg 5460
aaacacggaa accgaagacc attcatgttg ttgctcaggt cgcagacggt ttgcagcagc 5520
55 agtcgcttca cgttcgctcg cgtatcggtg attcattctg ctaaccagta aggcaacccc 5580
gccagcctag ccgggtcctc aacgacagga gcacgatcat gcgcacccgt ggccaggacc 5640

caacgctgcc cgagatgcgc cgcgtgcggc tgctggagat ggccggacgcg atggatatgt 5700
 tctgccaagg gttggtttgc gcattcacag ttctccgcaa gaattgattg gctccaattc 5760
 5 ttggagtggg gaatccgtta gcgaggtgcc gccggcttcc attcaggtcg aggtggccccg 5820
 gctccatgca ccgcgacgca acgcggggag gcagacaagg tatagggcgg cgcctacaat 5880
 10 ccatgccaac ccgttccatg tgctcgccga ggccggcataa atcgccgtga cgatcagcgg 5940
 tccagtgatc gaagttaggc tggtaagagc cgcgagcgat ccttgaagct gtccctgatg 6000
 gtcgtcatct acctgcctgg acagcatggc ctgcaacgcg ggcatcccga tgccgcccga 6060
 15 agcgagaaga atcataatgg ggaaggccat ccagcctcgc gtcgcgaacg ccagcaagac 6120
 gtagcccagc gcgtcggccg ccatgccggc gataatggcc tgcttctcgc cgaaacgttt 6180
 20 ggtggcggga ccagtgcga aggcttgagc gagggcgtgc aagattccga ataccgcaag 6240
 cgacaggccg atcatcgtcg cgctccagcg aaagcgggcc tcgccgaaaa tgaccagag 6300
 cgctgccggc acctgtccta cgagttgcat gataaagaag acagtcataa gtgcggcgac 6360
 25 gatagtcatg ccccgcgccc accggaagga gctgactggg ttgaaggctc tcaagggcat 6420
 cggtcgacgc tctcccttat ggcactcctg cattaggaag cagcccagta gtaggttgag 6480
 gccgttgagc accgcccgg caaggaatgg tgcagtctcg atggctacga gggcagacag 6540
 30 taagtggatt taccataatc ccttaattgt acgcaccgct aaaacgcgtt cagcgcgatc 6600
 acggcagcag acaggtaaaa atggcaacaa accaccctaa aaactgcgcg atcgcgccctg 6660
 35 ataaatttta accgtatgaa tacctatgca accagagggg acaggccaca ttacccccac 6720
 ttaatccact gaagctgcc ttttcatgg tttcaccatc ccagcgaagg gccatgcatg 6780
 catcgaaatt aatacgacga aattaatacg actcactata gggcaattgc gatcaccaca 6840
 attcagcaaa ttgtgaacat catcacgttc atctttccct ggttgccaat ggcccatttt 6900
 cctgtcagta acgagaaggt cgccaattca ggccgttttt agactggctg
 taatgaac 6958